Intraday FX Trading: An Evolutionary Reinforcement Learning Approach

Conference Paper *in* Lecture Notes in Computer Science · November 2002

DOI: 10.1007/3-540-45675-9_52

Michael A. H. Dempster
University of Cambridge

Civitas Foundation Finance Seminar

Princeton University, 20 November 2002

Intraday FX Trading: An Evolutionary Reinforcement Learning Approach

M A H Dempster

Centre for Financial Research
Judge Institute of Management
University of Cambridge
mahd2@cam.ac.uk – http://www-cfr.jims.cam.ac.uk

Co-worker: Y S Romahi

Outline

1. **Motivation**
2. **Problem Definition**
3. **Computational learning techniques**
   - Reinforcement Learning
   - Evolutionary Learning
   - Evolutionary Reinforcement Learning
4. **Results**
   - Performance of the three systems
   - What are the implications of different frequency trading?
   - Should the indicators fed into the system be pre-optimized?
   - Does allowing a neutral state improve or hinder performance?
5. **Conclusions**

Motivation

- Increasing evidence that markets are predictable
  - Lo & McKinley state that rather than being a symptom of inefficiency, predictability in the financial markets is the "oil that lubricates the gears of capitalism"
- Most technical traders are active in the FX markets and at high frequency
  - Daily *vs* high frequency [Neeley (1999)]
  - Equities *vs* FX [Taylor & Allen (1992)]
  - Asset allocation *vs* trading [Dempster & Jones (2001)]

Previous Work

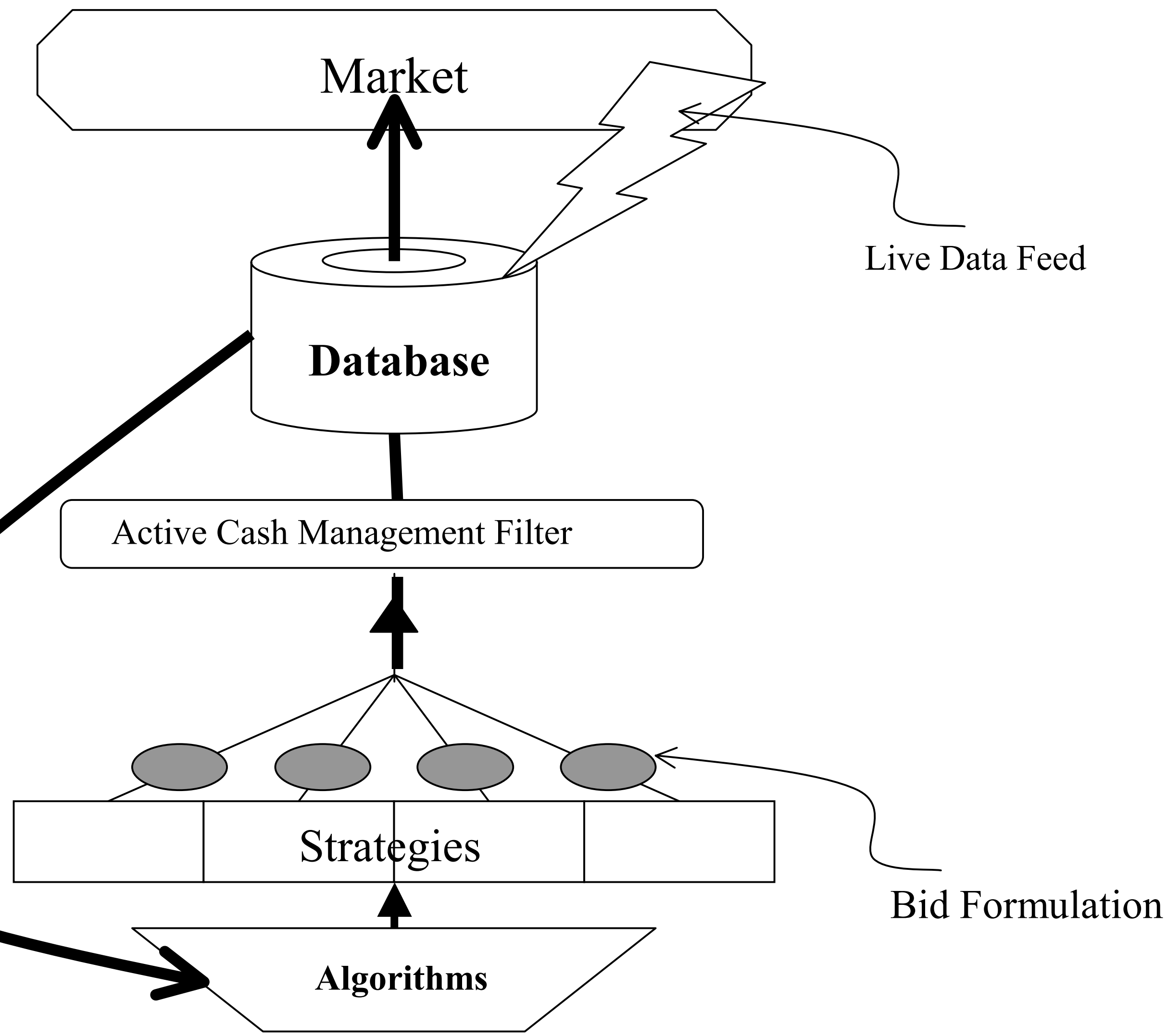
- Previous work examining popular indicators does not find evidence of profit opportunities, e.g. [Dempster & Jones (1999)]
- [Neeley, Weller and Dittmar (1997)] found out-of-sample annual excess returns in the 1-7% range in currency markets against the dollar during 1981-1995
- [Dempster et al. (2001)] found significant out-of-sample annual returns up to about 2bp slippage using various computational learning methods

Intelligent Trading Systems

- Trading System
  - Trading systems generate **buy rules**, **sell rules** and **exit rules**
  - Rules are defined as a mapping between **states** and **actions**
  - States are defined as a combination of **indicators** (which can be technical/fundamental/composite)

- Key Features of an Intelligent Trading System
  - **Learning & Discovery**
  - **Adaptation**
  - **Explanation**

The System in a 'Live' Trading Context



Adaptive Trading

- Strategies are **combinations of technical indicators** drawn from the worlds of technical analysis
- Best performing strategies are selected using **computational learning techniques**
- System can be overlaid by a simple **cash/risk management filter**
- **Adaptation** is achieved in several ways:
  - Online Learning
  - Re-mining at set intervals
  - Profitability dependent time intervals
    - (e.g. if portfolio loses money for *n* consecutive periods)

Data

- 4 years of high frequency (1 minute) midpoint data from Bank of America
- Currencies: GBPUSD, USDCHF and USDJPY
- Frequencies: 15 minute, 1 hour, 2 hour, 4 hour and 8 hour
- Dates: January 1995 to December 1998 (1995 used for first in-sample learning period)

The Need For High Frequency Data

- To simulate the actions of traders
  - ‘desk traders’ watch the markets ‘tick by tick’ and apply the concepts of technical analysis at frequencies much higher than ‘daily’
- Trade entry and exit strategies
  - Even technical traders who look for patterns in daily data alone often use tick data for confirmatory entry signals
  - The vast majority of traders place stops in the markets alongside their trades to manage downside risk and such stops are activated at ‘tick’ level
- In general - realism

Objective Function

- Simulate simple trader in single currency pair
  - Trades by drawing on a credit line, converting, holding and then converting back and accumulating any profit/shortfall in domestic currency (dollars)
  - Can borrow \$1 (or equivalent) in either currency
  - Cumulated profit or loss at end of sample period is objective value
- Transaction costs (due to bid-ask spread and slippage) charged at 0, 1, 2 and 4 basis points of amount exchanged

More Formally…

- With transaction cost c exchange rates (expressed per unit of home currency) of F_t at trade entry and $F_{t'}$ at trade exit drawing on a credit line of C units of home currency and taking a long position in the foreign currency will yield a return per unit of home currency of

$$\left[\frac{F_t}{F_{t'}}(1-c)^2 - 1\right]$$

- If a short position is taken in the foreign currency then C/F_t units of foreign currency are drawn from the credit line and the return per unit of home currency is:

$$\left[(1-c) - \frac{F_{t'}}{F_t}\frac{1}{(1-c)}\right]$$

Objective Function

- Indicator signals over time **s** a *stochastic process* with state space ***S*** driven by the exchange rate process **F**
- Solve the *stochastic optimization problem* defined by the maximisation of expected return over the trading horizon net of transaction costs

$$\mathbb{E} \sum_{i=1}^{\mathbf{N}_T} r_i \left(F_{t_i}, F_{t_i'} \right)$$

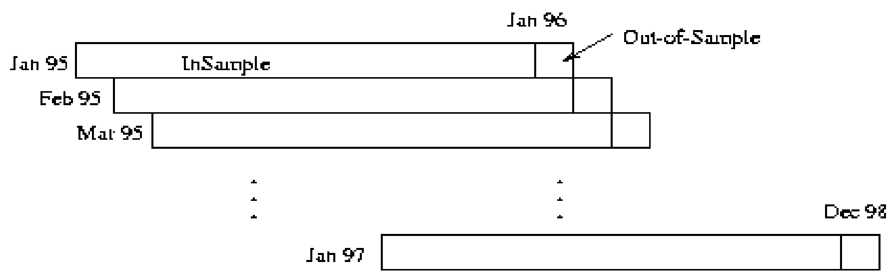
- The statistics of the processes **F** and **s** are entirely unknown
- Computational learning methods attempt to find approximate solutions by discovering a (feedback) *trading strategy* ϕ: $\mathcal{S}$ x {l,s} $\rightarrow$ {l,s} that maps the current market state $\mathbf{s}_t$ and position to a new position

Technical Indicators

- A number of popular indicators used as inputs:
  - Momentum Oscillator
  - Price Channel Breakout
  - Moving Average Crossover
  - Relative Strength Index
  - Adaptive Moving Average
  - Moving Average Convergence/Divergence
  - Stochastics
  - Commodity Channel Index

Problem Definition

- Technical indicators together define market state
- System attempts to learn what trading action to take in each state. Two systems are examined;
  - **2-way system**: Always in the market (long/short positions)
  - **3-way system**: Can take neutral positions (out of the market)
- Train on 12-month moving window, use 1 month out-of-sample



Genetic Algorithm

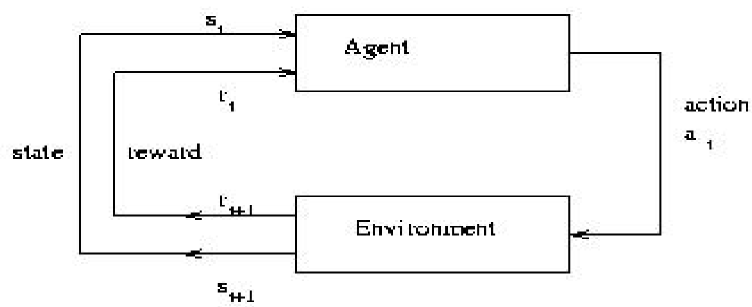
- GA evolves trading rules
- Rules are represented as binary trees
  - Terminals are indicators (MO, PCB, MAX, etc.)
  - Non-terminals are Boolean functions (AND, OR, XOR)
  - Together represent simple functions (e.g. IF MO AND PCB THEN LONG)
- Rule is executed at each time step to get position at next time step
- Best in-sample performing rule (on raw return) is used out-of-sample

Reinforcement Learning

- Attempts to discover which trading action to take in each market state (as defined by indicators) by receiving rewards; aim is to maximise total reward
- Reward function is increase in wealth
- Makes multiple passes through in-sample period, refining its actions each time
- Uses Watkins' Q-Learning algorithm so rewards 'trickle down' to actions that caused them

Reinforcement Learning

- Learning while interacting with the environment

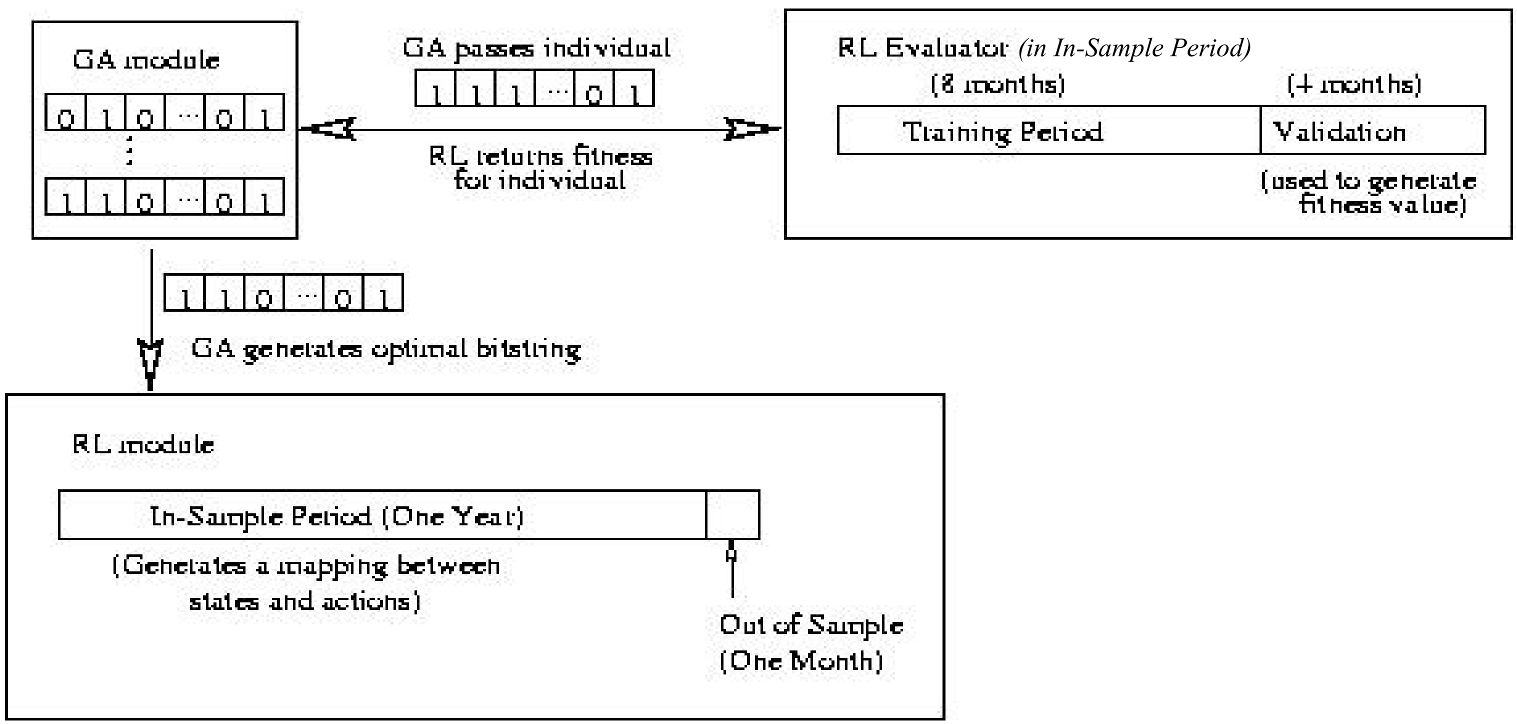


- RL methods estimate value functions (defined in terms of total reward) based upon previously learned estimates

$$Q(s_t, a_t) \Leftarrow Q(s_t, a_t) + \alpha \left[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t - a_t) \right]$$

Evolutionary RL

- RL was prone to overfitting when given too many inputs [Dempster et al. (2001)]
- Can we use a Genetic Algorithm to constrain the inputs to RL?

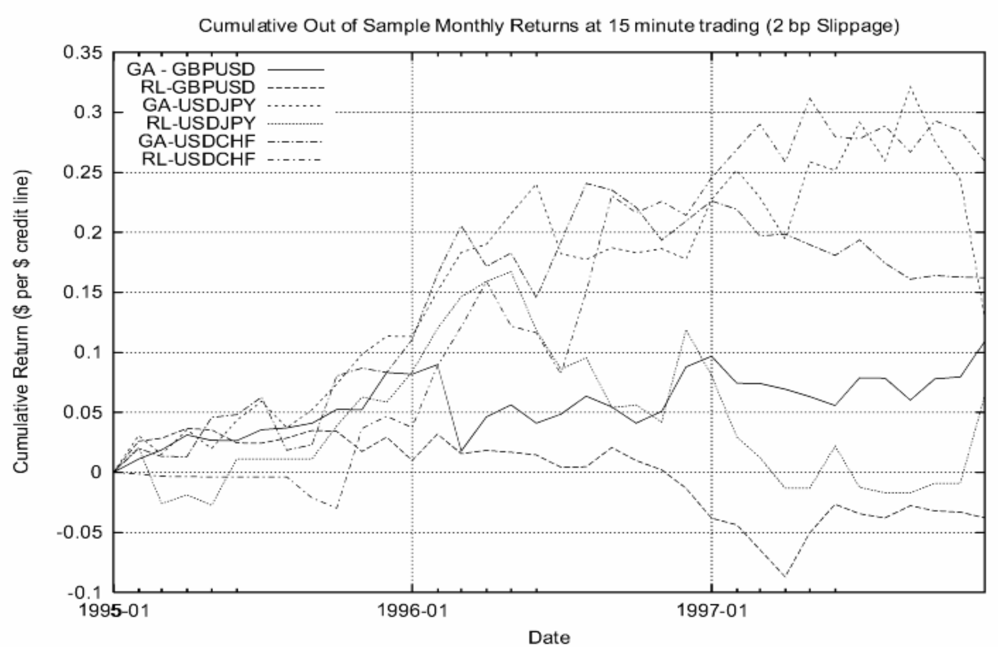


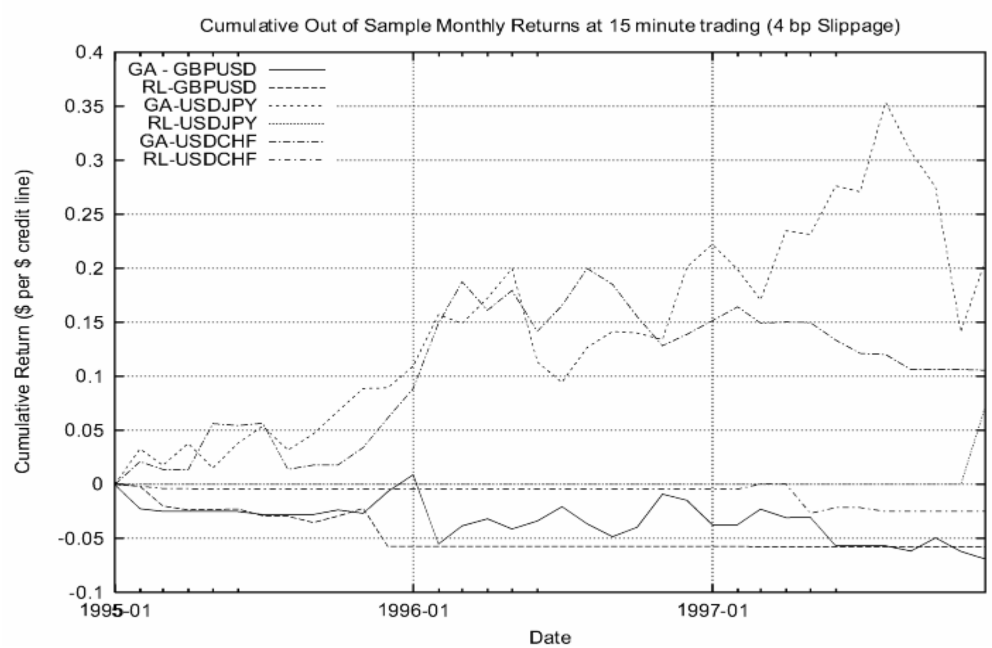
Evolutionary RL

- Genetic Algorithm bitstring defines the indicators that are fed into the RL

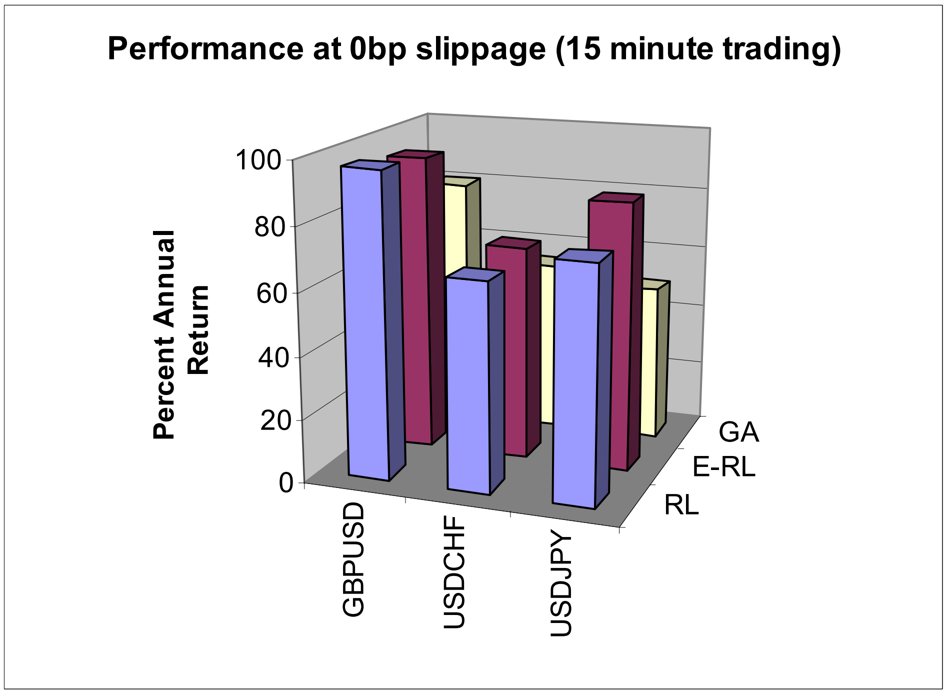
- The fitness function is defined as maximum return or some measure of risk adjusted return

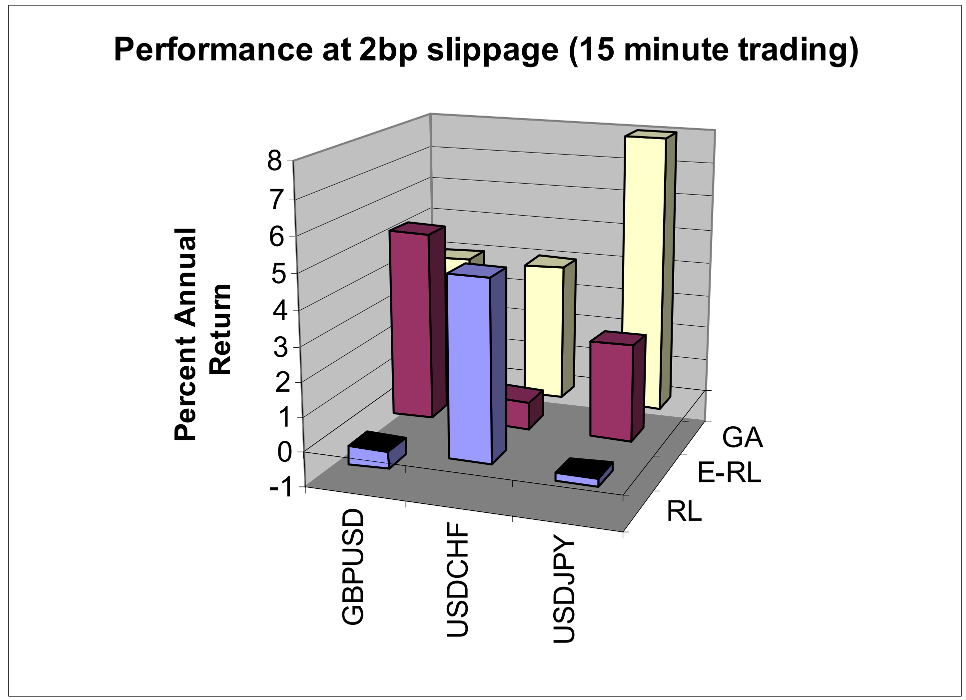
Reinforcement *vs* Evolutionary Learning



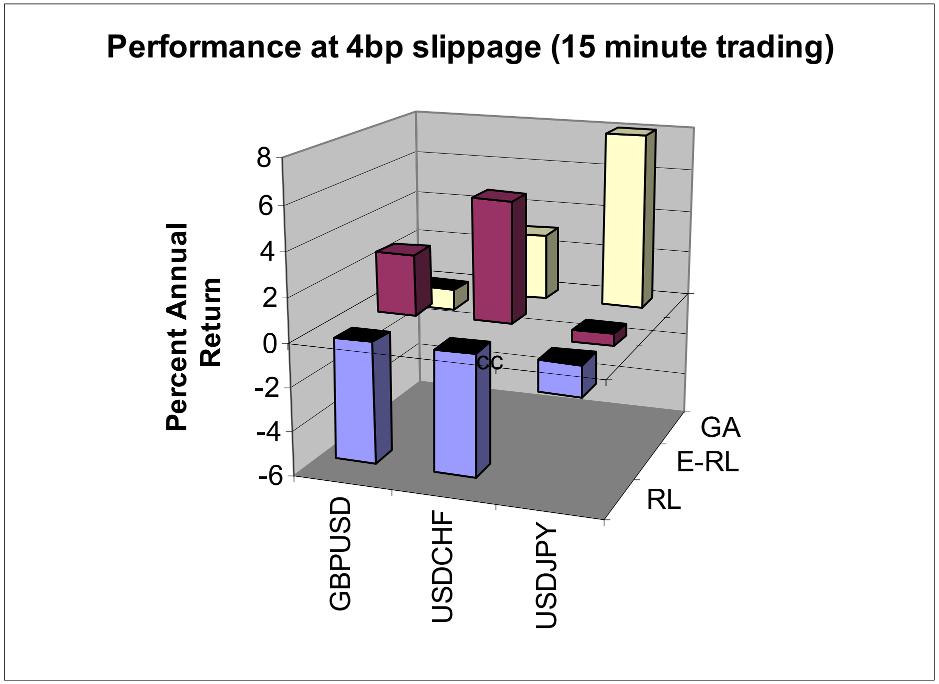


Results: Comparison of Methods (Two State)





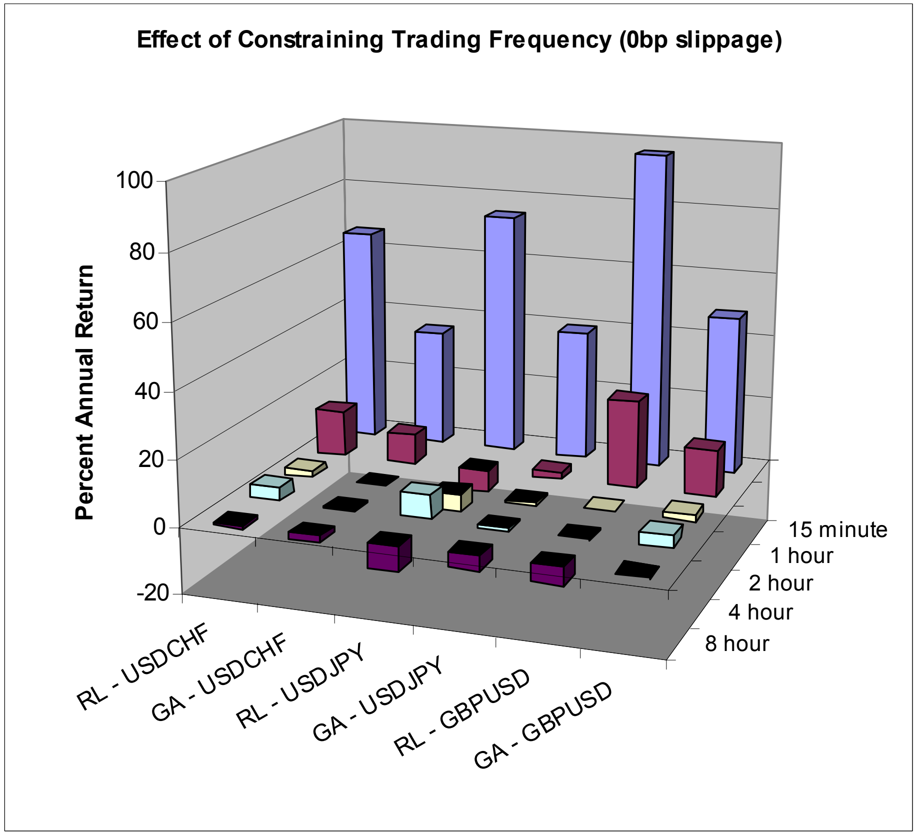
Results: Comparison of Methods (Two State)

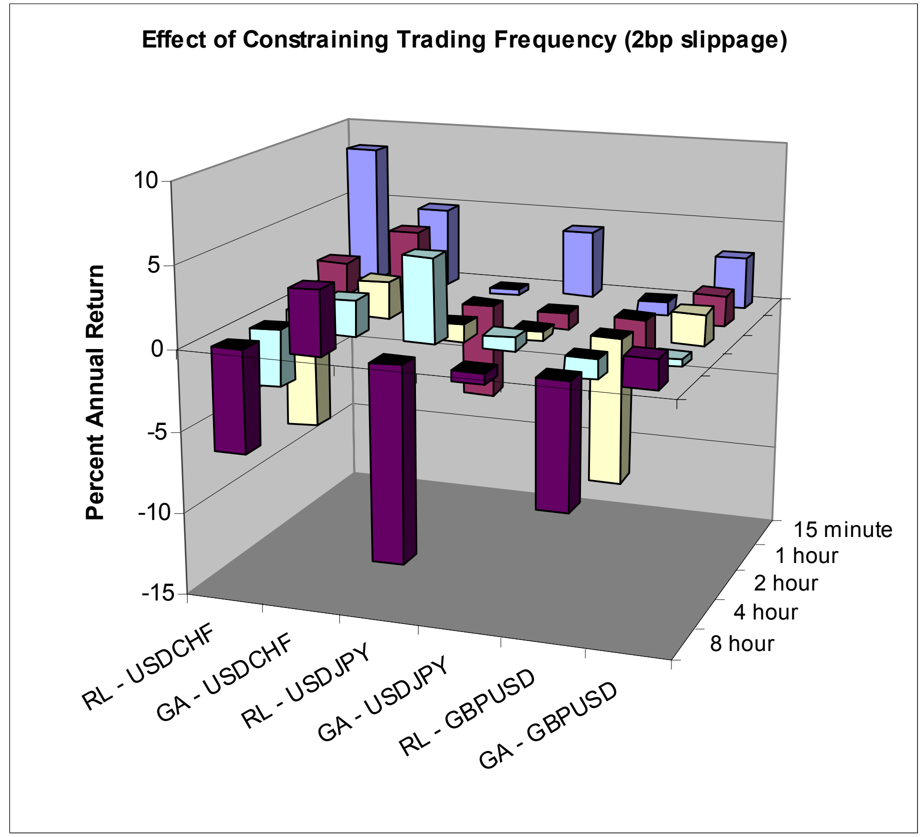


Results: Comparison of Methods

- Returns are broadly similar at 0bp & 1bp slippage
- At higher slippage values GA is able to trade profitably at up to 4bp whereas RL is unable to trade profitably beyond 1bp
- RL suffers from overfitting the in-sample data
- ERL system overcomes this issue and obtains the highest returns at 4bp in 2 of the 3 currencies studied

Constraining Trading Frequency





Constraining Trading Frequency

RL – Average Monthly Trading Frequency					
Slippage	**0 bp**	**1 bp**	**2 bp**	**4 bp**	**8 bp**
GBPUSD – 15 minute	324.52	135.14	112.44	0.36	0
GBPUSD – 1 hour	71.61	37.78	29.53	4.64	0.39
GBPUSD – 2 hour	38.61	22.33	8.97	3.17	2.58
GBPUSD – 4 hour	16.31	11.92	10.56	7.22	4.61
GBPUSD – 8 hour	11.19	8.17	6.92	6.14	1.89

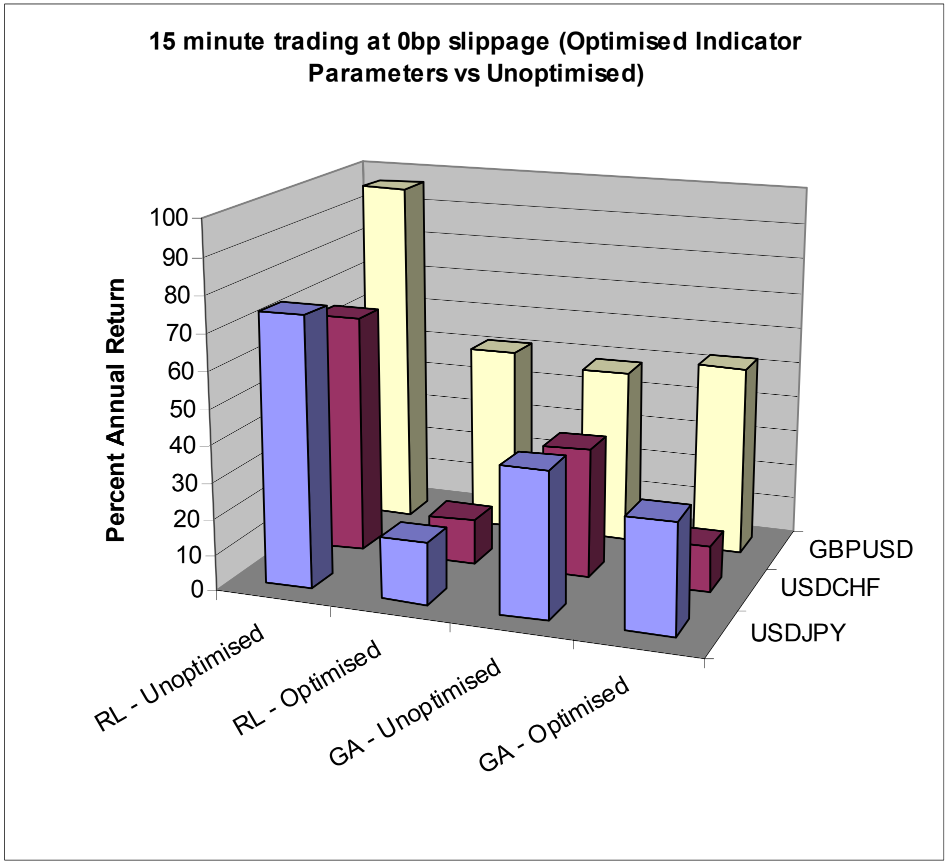
RL – Average Monthly Trading Frequency					
Slippage	**0 bp**	**1 bp**	**2 bp**	**4 bp**	**8 bp**
USDJPY – 15 minute	492.25	86.81	58.53	5.69	0.39
USDJPY – 1 hour	75.86	31.11	13.30	4.67	3.33
USDJPY – 2 hour	37.42	22.11	17.31	15.61	9.67
USDJPY – 4 hour	21.39	11.72	10.13	8.56	4.83
USDJPY – 8 hour	11.03	9.19	8.67	7.78	5.25

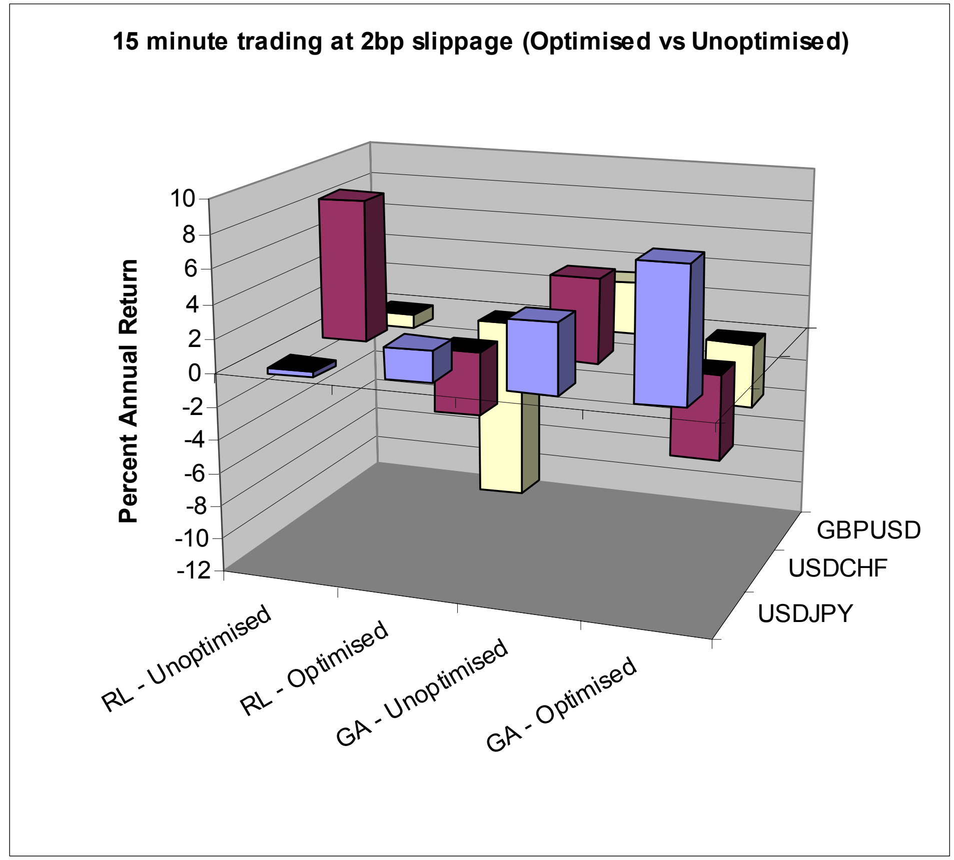
Note that results for GA were broadly similar

Frequency Effect

- As slippage increases monthly dealing adapts to frequency
- Rather than constrain the frequency artificially in the pre-processing step, the algorithm should be left to choose its own trading frequency by feeding it with the highest frequency data available

Optimized *vs* Unoptimized Performance

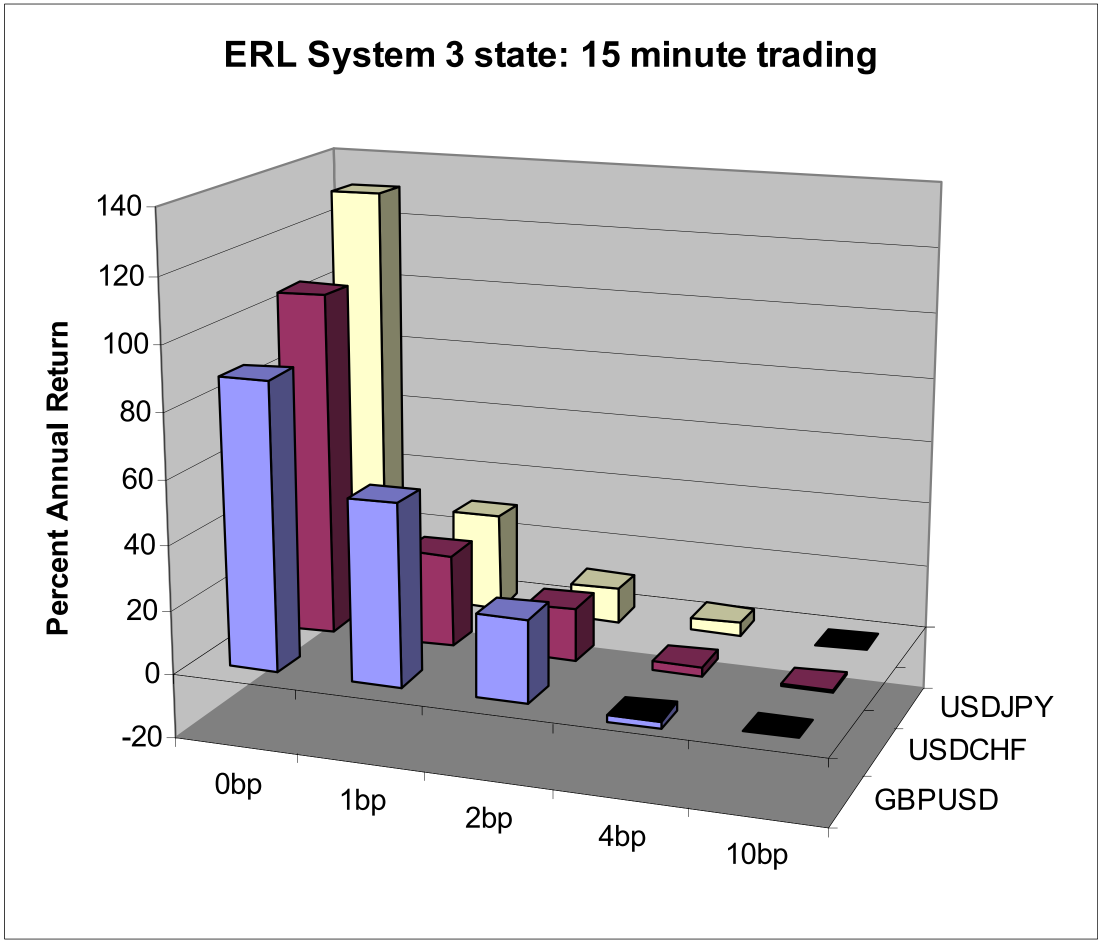




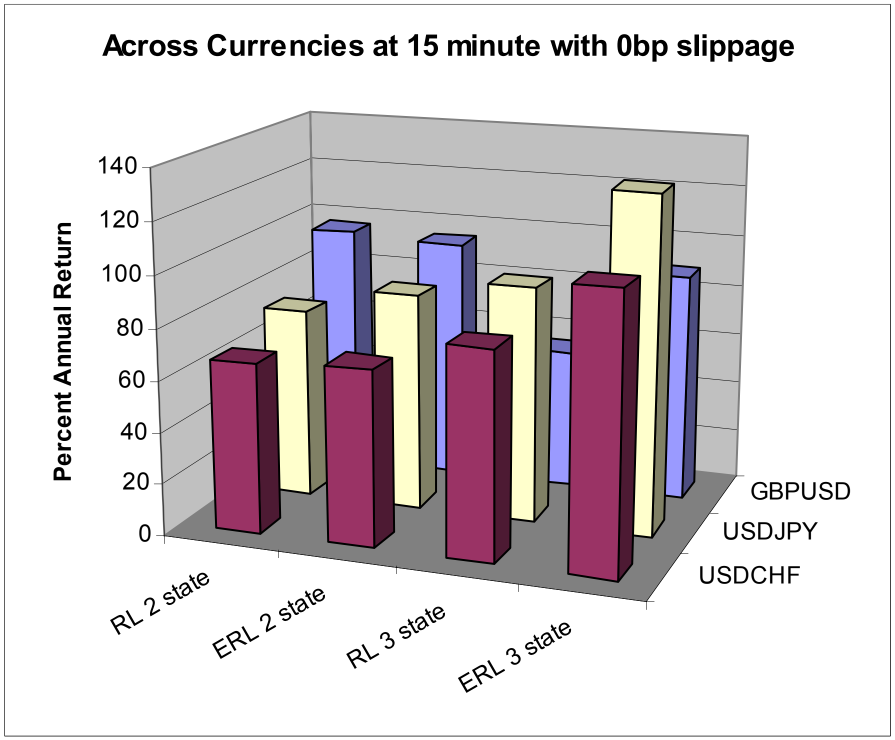
Should We Optimize Parameters?

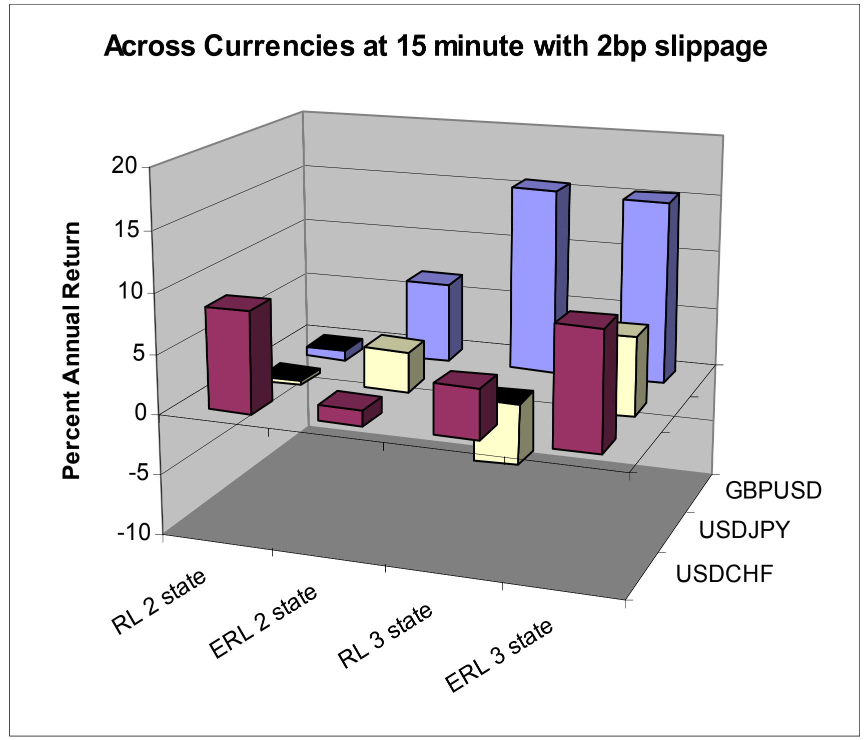
- Indicators attempt to capture similar dynamics
- By optimizing indicator parameters the correlation between the indicators increases significantly
- Thus information content for the learning methods decreases

Results: Evolutionary RL System 3-state



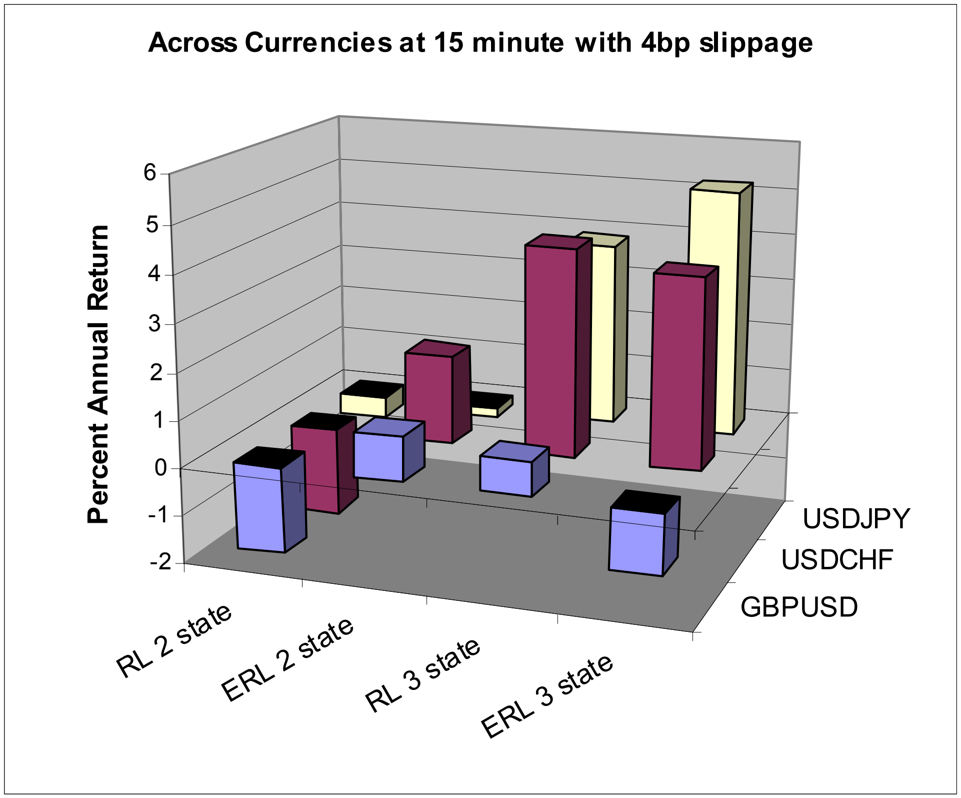
Results: 2-state vs 3-state





- We examine whether allowing a neutral state improves the system or impedes performance

Results: 2-state vs 3-state



- These results demonstrate the advantages of allowing a neutral state as the 3-state system consistently outperforms its 2 state counterpart

Significance Test

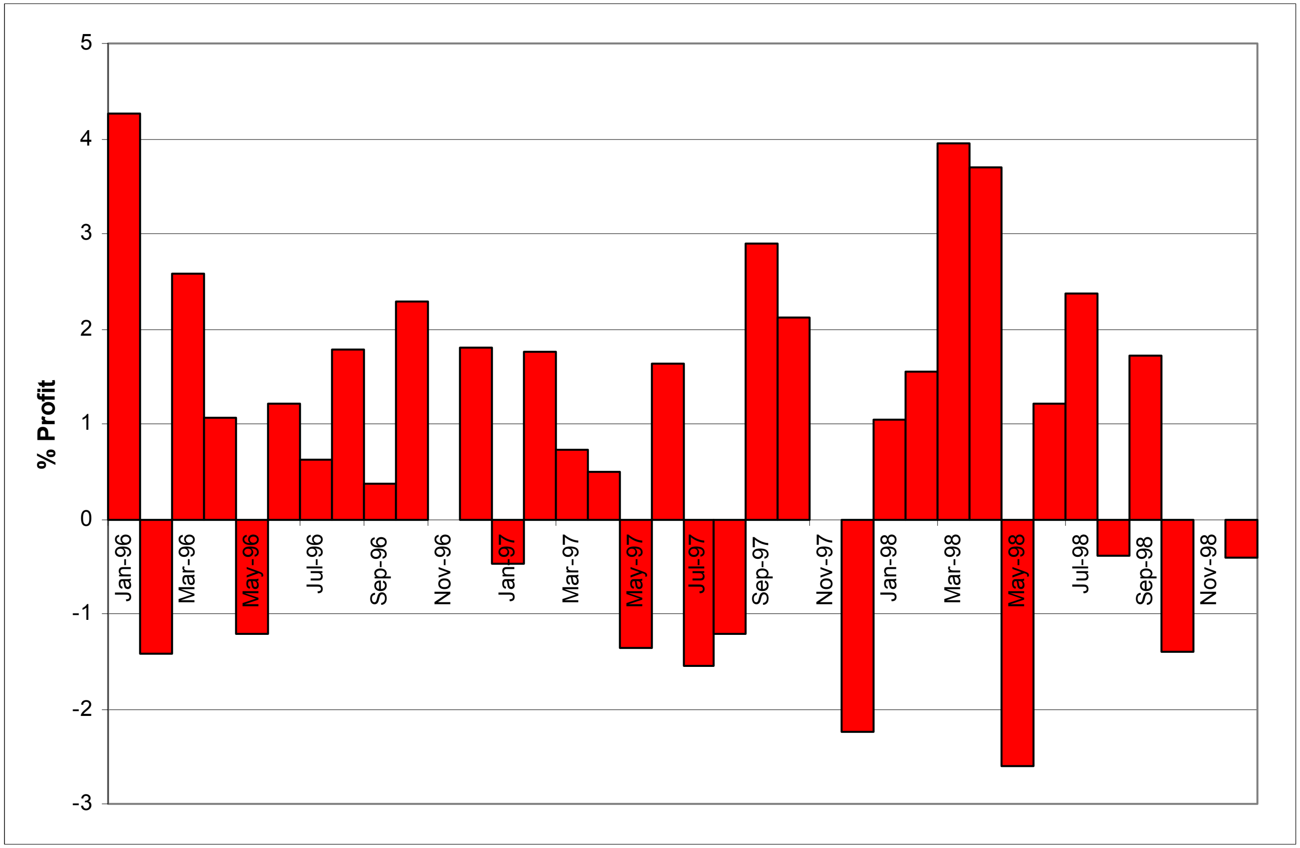
- We utilize a simple non-parametric binomial test [Dempster and Jones (2001)]
- Null hypothesis : out-of-sample cumulative trading profits and losses are periodically sampled from a continuous time stationary ergodic process with state distribution having median zero
- Under this null hypothesis, profits and losses are equally likely
- It follows that over n monthly periods, the number of profitable months n+ is binomially distributed with parameters n and ½
- We test the two-tailed hypothesis that median profit and loss is non-zero with the statistic n+

Significance of Two-State Results

p-value			
Trading Frequency	GBPUSD	USDCHF	USDJPY
RL 0bp	**100**	**99.99**	**99.99**
GA 0bp	**100**	**99.99**	**99.99**
ERL 0bp	**100**	**100**	**100**
RL 1bp	**100**	**95.22**	**90.98**
GA 1bp	**99.99**	**95.22**	**99.02**
ERL 1bp	**100**	**99.99**	**95.22**
RL 2bp	9.12	25.25	63.06
GA 2bp	**90.88**	50	84.12
ERL 2bp	**95.22**	63.06	84.12
RL 4bp	50	15.86	36.94
GA 4bp	25.25	74.75	84.13
ERL 4bp	74.75	63.06	50

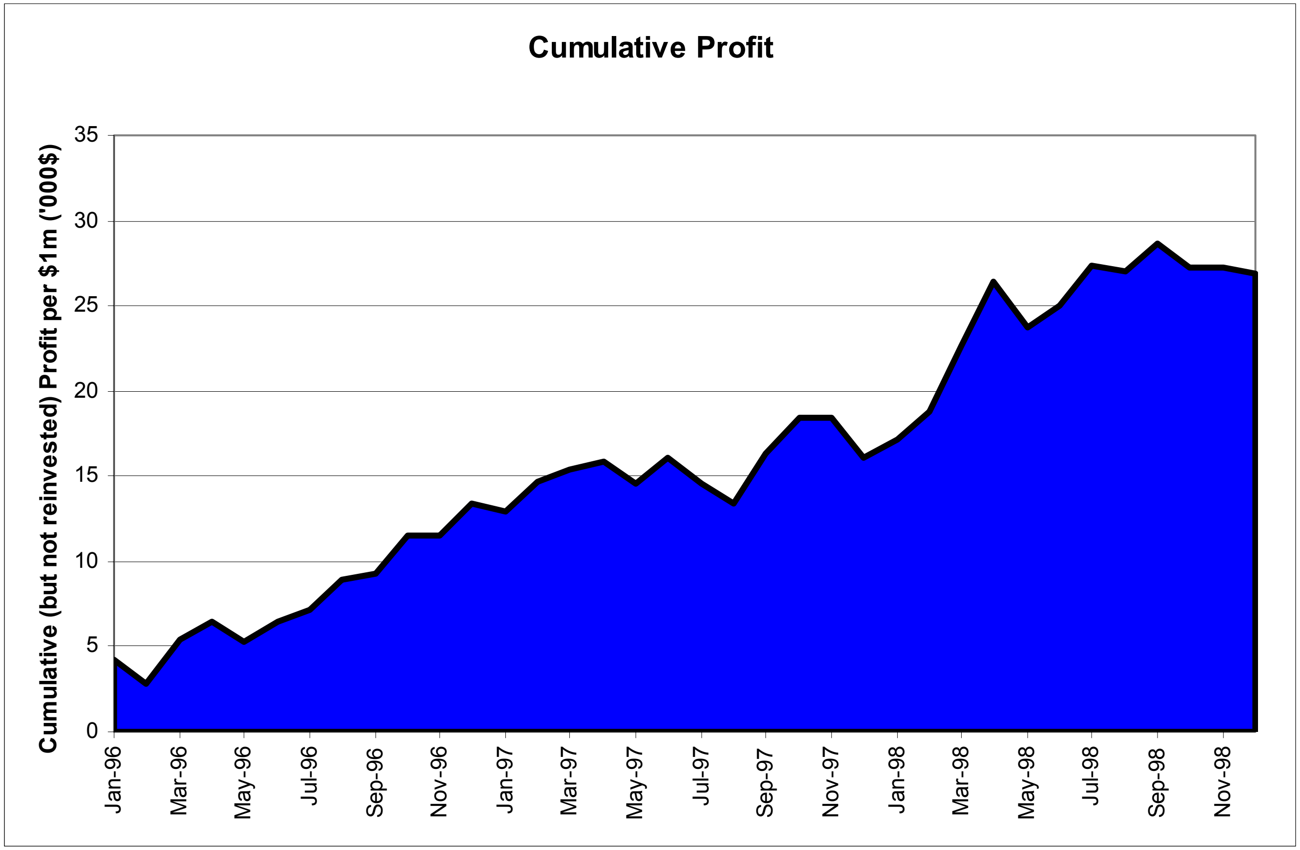
At 8bp and 10bp the results were uniformly not significant

Evolutionary RL System – USDCHF 2-state 15 minute at 2bp



p-value 0.9082

Evolutionary RL System – USDCHF 2-state 15 minute at 2bp

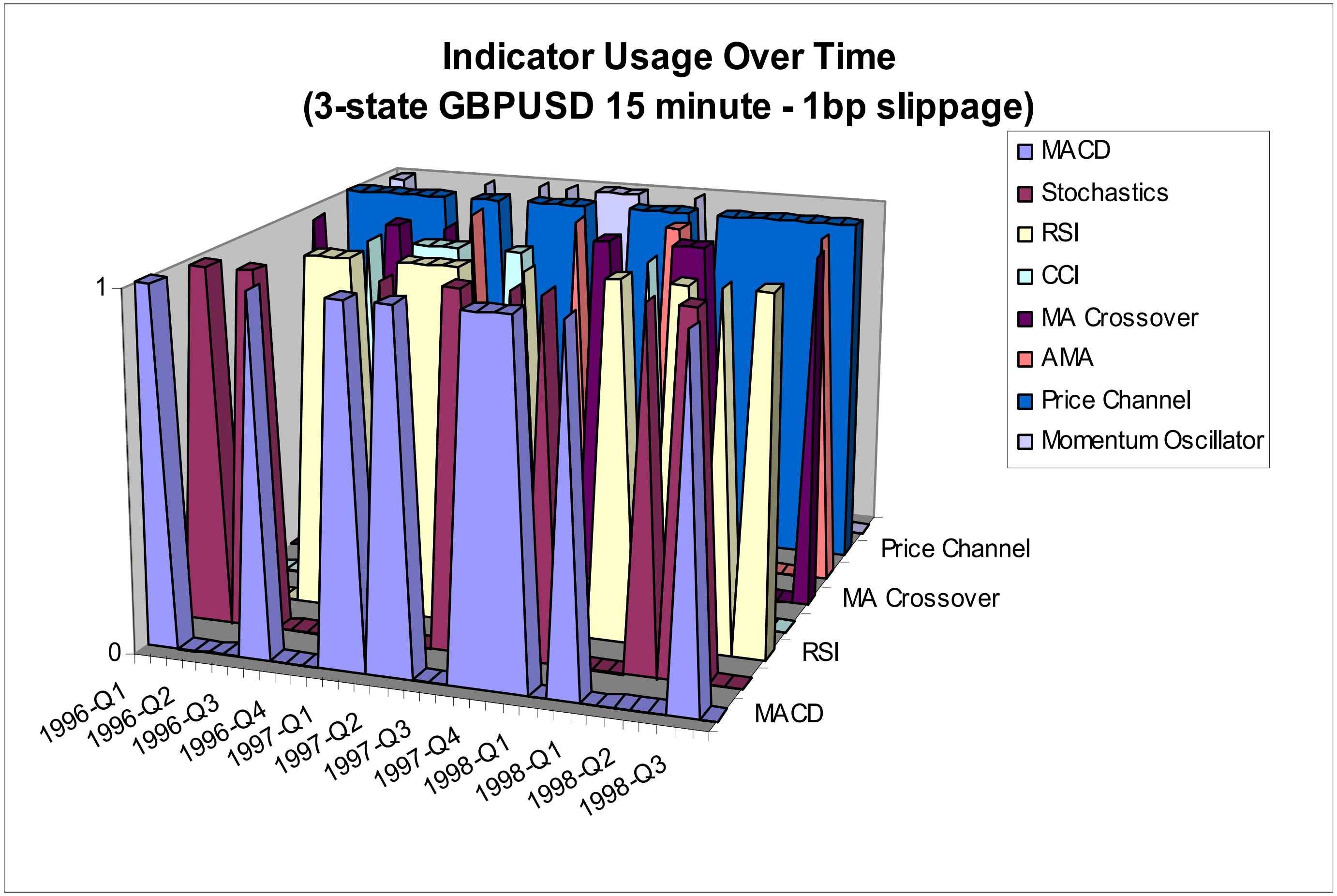


Significance of Three State Result

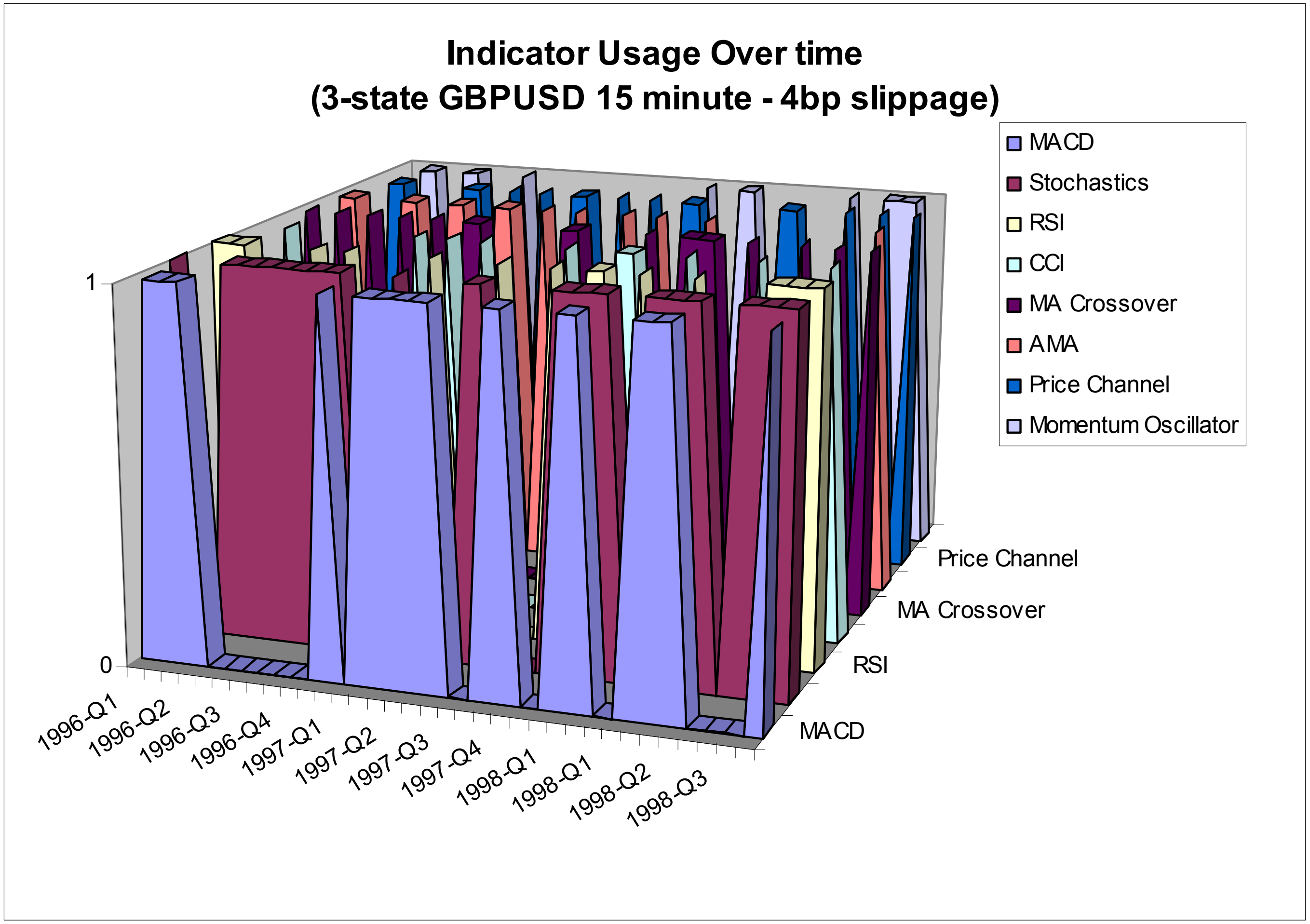
p-values			
Trading Frequency	GBPUSD	USDCHF	USDJPY
RL 0bp	**100**	**100**	**100**
ERL 0bp	**100**	**100**	**100**
RL 1bp	**100**	**95.22**	**95.22**
ERL 1bp	**99.99**	**99.99**	**99.99**
RL 2bp	**99.02**	**90.88**	65.54
ERL 2bp	**99.99**	**99.02**	**95.22**
RL 4bp	63.06	80.75	84.13
ERL 4bp	**90.88**	**90.88**	**95.22**

At 8bp and 10bp the results were uniformly not significant

Indicator Usage by the ERL System



Indicator Usage by the ERL System



Indicator Usage Over Time

- The systems **continuously adapt** as market dynamics change and the indicators chosen do not remain static
- Certain indicators are consistently favoured by the evolutionary subsystem at low slippage (0bp and 1bp) while different indicators are favoured at the higher (4bp) slippage value

Optimizing Risk Adjusted Return

- RL typically requires immediate rewards
  - RL runs into difficulty associating delayed negative feedback with the states that caused them
  - Difficult to incorporate drawdown or Sharpe ratio as these are not *instantaneous* measures

- Evolutionary RL system is ideally suited to solve this
  - Introduction of risk adjusted optimization moved to the GA layer where it is straightforward to incorporate
  - GA wrapper fitness function maximizes some measure of risk-adjusted reward over the evaluation period
  - RL optimization remains as discussed earlier

Results: Risk Adjusted Returns

- Results:
  - The results demonstrated that the risk parameters improved although returns remained broadly in line with those outlined previously
  - The quoted results illustrate the Sharpe Ratios of two ERL systems –the optimization of a drawdown adjusted return compared to total return

Sharpe Ratios	GBPUSD	USDCHF	USDJPY
Drawdown Adjusted Return (0bp)	**2.33**	**2.3**	**1.84**
No Risk Adjustment (0bp)	2.21	2.22	1.82
Drawdown Adjusted Return (1bp)	**1.46**	**0.52**	**0.41**
No Risk Adjustment (1bp)	1.51	0.5	0.37
Drawdown Adjusted Return (2bp)	**0.79**	**0.4**	**0.2**
No Risk Adjustment (2bp)	0.7	0.25	0.13
Drawdown Adjusted Return (4bp)	**0.02**	**0.24**	**0.07**
No Risk Adjustment (4bp)	-0.04	0.16	0.12

Conclusions

- RL *vs* EL and ERL
  - RL tends to over-fit
  - Need to constrain in-sample learning
  - This is what was done with the ERL approach
  - ERL was shown successfully to improve out-of-sample performance
  - GA is *significantly* profitable at up to 2bp
  - GA is able to trade profitably up to 4bp
  - GA and hybrid ERL results similar and allow for easier analysis of the indicators being used

Conclusions 2

- Frequency effect
  - Trading systems should be left to adapt to the frequency automatically. This is done based on the slippage – rather than artificially constraining the inputs in the pre-processing step
- Indicator optimization
  - optimization does not appear to improve results due to increased correlation amongst indicators
- Across currencies
  - Results across currencies were broadly in line giving further confidence in the results

Further Work

- Hybrid ERL system has shown to be an improvement over the basic RL and competitive with the GA

- Areas currently being studied:
  - Optimizing risk adjusted return
  - Overlay of cash management strategies
  - Alternative forms of reinforcement learning
  - Incorporation of trade volume and order flow data

Data

- 5 ¾ months of order book and order flow data
- Currencies: GBPUSD, EURUSD and USDJPY
  - USDJPY was not considered for the order book due to errors in the file
- Intraday order book/flow data but for this pilot study was collated into **daily**
- Dates: March 1, 2002 – August 19, 2002
- Insample period: 1st March – 30th May (65 data pts)
- Out-of-sample: 1st of June – 19th August (58 data pts)

Order Flow Data

- HSBC broke down transactions into several categories
- These categories were used to derive the indicators:
  - **Net daily** transactions in that currency pair (1 if positive volume, 0 if negative)
  - **Net retail** transactions
  - **Retail speculative** transactions
  - **Retail nonspeculative** transactions
  - **Net institutional** transactions
  - **Institutional speculative** transactions
  - **Institutional nonspeculative** transactions
  - **Net speculative** Transactions
  - **Net nonspeculative** transactions

Order Book Data

- For each day the following **indicators** are generated:
  - Net customer sales for stop-loss orders where the price is more than 0.0% and less than or equal to 0.5% from the current spot.
  - Net customer sales for stop-loss orders where the price is more than 0.5% and less than 1% from the current spot.
  - Net customer sales for stop loss orders where the price is between 0 and 1% (i.e. the sum of the former two)
- These are calculated for **all** orders and for **take-profit** orders for the **whole** order book as at the time of the snapshot of the book and for **new** orders only
- The total number of indicators derived from the order book is 12

Order Book Data Return Correlations

- Interestingly correlations are significant at a lag of 2 days
- Similar indicators are found to be significant in the two currencies
- Indicators of stop losses at more than 0.5% of the spot appear to be most significant

GBPUSD Net Customer Sales	FXReturns	lag(1)	lag(2)
New: 0%-0.5% of spot	4.572 (0.46)	-0.292 (-0.02)	5.114 (0.51)
New: 0.5%-1% of spot	6.895 (0.69)	-4.729 (-0.47)	-9.305 (-0.93)
New: 0%-1% of spot	8.37 (0.84)	-3.888 (-0.39)	-3.943 (-0.39)
New: 0%-0.5% of spot (TP)	7.643 (0.77)	-1.188 (-0.11)	7.455 (0.74)
New: 0.5%-1% of spot (TP)	-12.445 (-1.26)	-3.015 (-0.3)	**17.744 (1.8)**
New: 0%-1% of spot (TP)	-3.17 (-0.32)	-3.258 (-0.32)	***19.549 (1.99)***
All: 0%-0.5% of spot	2.102 (0.21)	5.597 (0.56)	1.481 (0.14)
All: 0.5%-1% of spot	4.675 (0.47)	-4.269 (-0.42)	**-17.329 (-1.75)**
All: 0%-1% of spot	4.5 (0.45)	-0.23 (-0.02)	-11.896 (-1.19)
All: 0%-0.5% of spot (TP)	4.395 (0.44)	1.943 (0.19)	1.313 (0.13)
All: 0.5%-1% of spot (TP)	2.597 (0.26)	-0.084 (0)	9.262 (0.93)
All: 0%-1% (TP)	5.768 (0.58)	1.498 (0.15)	8.96 (0.89)

EURUSD Net Customer Sales	FXReturns	lag(1)	lag(2)
New: 0%-0.5% of spot	13.502 (1.38)	14.2 (1.44)	-5.17 (-0.52)
New: 0.5%-1% of spot	10.729 (1.09)	9.209 (0.93)	-13.178 (-1.33)
New: 0%-1% of spot	***19.407 (2)***	**18.871 (1.94)**	-14.16 (-1.43)
New: 0%-0.5% of spot (TP)	3.411 (0.34)	10.077 (1.02)	-3.62 (-0.36)
New: 0.5%-1% of spot (TP)	-4.102 (-0.41)	0.335 (0.03)	***24.358 (2.52)***
New: 0%-1% of spot (TP)	-0.988 (-0.1)	7.07 (0.71)	**17.094 (1.74)**
All: 0%-0.5% of spot	12.573 (1.28)	8.606 (0.87)	-6.34 (-0.63)
All: 0.5%-1% of spot	5.752 (0.58)	0.21 (0.02)	***-19.556 (-2)***
All: 0%-1% of spot	15.201 (1.56)	7.624 (0.77)	***-20.098 (-2.06)***
All: 0%-0.5% of spot (TP)	2.652 (0.26)	12.175 (1.23)	-1.847 (-0.18)
All: 0.5%-1% of spot (TP)	-3.19 (-0.32)	3.992 (0.4)	***27.469 (2.87)***
All: 0%-1% (TP)	-0.383 (-0.03)	11.799 (1.2)	***18.633 (1.9)***

Preliminary Results – Benchmark (Technical Indicator) Tests

Currency	Slippage	In-sample	Out-of-sample
EURUSD	0	1.46	0.09
EURUSD	2	1.02	0.47
EURUSD	4	0	0
EURUSD	8	0	0
EURUSD	10	0	0
GBPUSD	0	1.14	0.82
GBPUSD	2	1.64	0.36
GBPUSD	4	1.72	2.02
GBPUSD	8	1.16	0.16
GBPUSD	10	0	0
USDJPY	0	0	0
USDJPY	2	3.23	1.54
USDJPY	4	0	0
USDJPY	8	0.14	0.94
USDJPY	10	0	0

Table 1: Average Monthly Returns on Benchmark Tests

Currency	Slippage	In-sample	Out-of-sample
EURUSD	0	19	9.82
EURUSD	2	21	11.27
EURUSD	4	0	0
EURUSD	8	0	0
EURUSD	10	0	0
GBPUSD	0	16	14.91
GBPUSD	2	2.33	1.82
GBPUSD	4	1.67	1.09
GBPUSD	8	1.67	1.82
GBPUSD	10	0	0
USDJPY	0	0	0
USDJPY	2	27.67	15.64
USDJPY	4	0	0
USDJPY	8	1.33	1.09
USDJPY	10	0	0

Table 2: Average Number of Monthly Trades: Benchmark Technical Tests

Preliminary Results – Technical/Order Book Tests

Currency	Slippage	In-sample	Out-of-sample
EURUSD	0	3.22	0.17
EURUSD	2	2.26	1.17
EURUSD	4	1.57	-0.75
EURUSD	8	3.66	1.52
EURUSD	10	2.38	0.81
GBPUSD	0	1.05	1.25
GBPUSD	2	0.97	1.68
GBPUSD	4	1.1	0.4
GBPUSD	8	0	0
GBPUSD	10	0	0

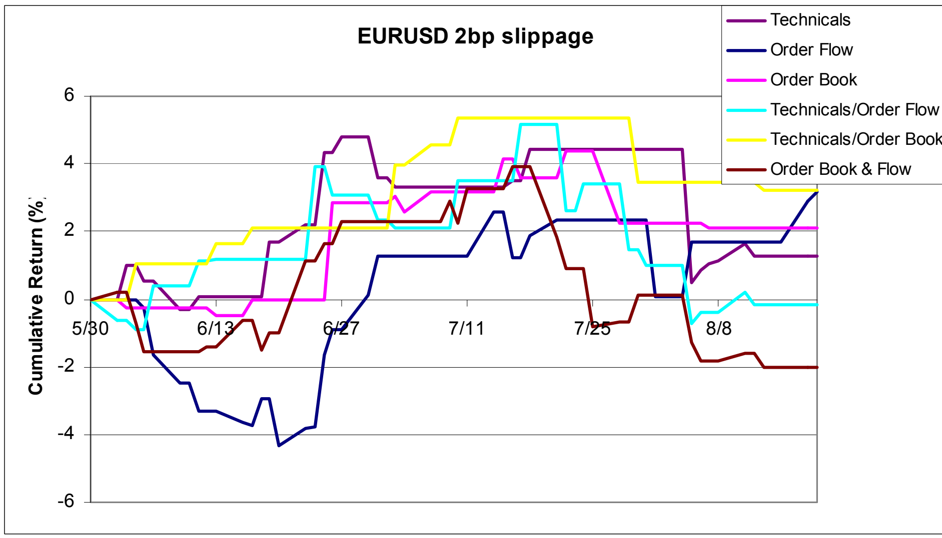
Table 9: Average Monthly Returns on Joint Technical and Order Book Tests

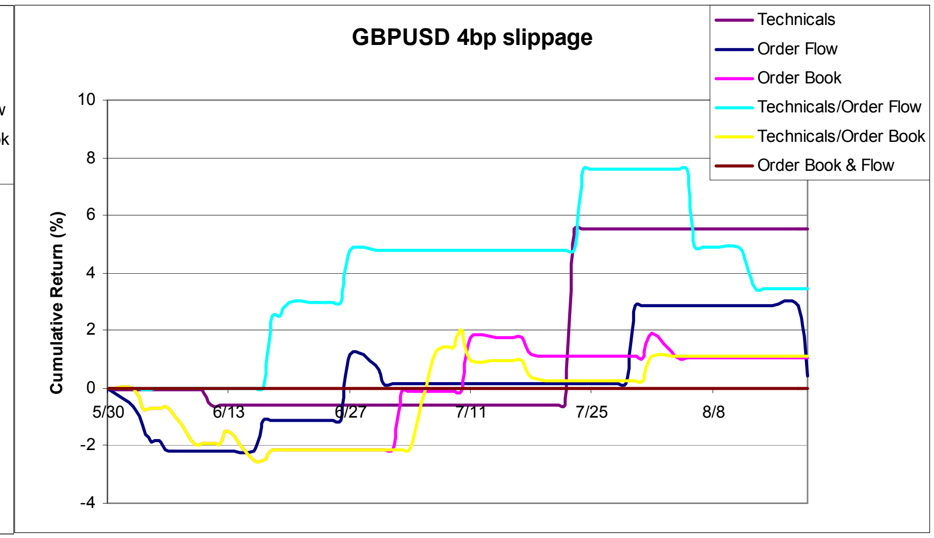
Currency	Slippage	In-sample	Out-of-sample
EURUSD	0	10	13.45
EURUSD	2	6	5.45
EURUSD	4	10.67	11.27
EURUSD	8	5.67	6.18
EURUSD	10	4	2.91
GBPUSD	0	4.67	5.82
GBPUSD	2	7.67	10.18
GBPUSD	4	4	5.09
GBPUSD	8	0	0
GBPUSD	10	0	0

Table 10: Average Number of Monthly Trades: Joint Technical and Order Flow Tests

- Merging technical with order flow indicators improved the out-of-sample returns
- Merging technical with order book indicators proved to be the best combination in both currencies examined

Performance over time





- To illustrate the cumulative return over the out-of-sample period, these are the graphs of EURUSD at 2bp slippage and GBPUSD at 4bp
- The report contains a full set of graphs for EURUSD and GBPUSD

Order Book Indicators Used by GA

Currency	Slippage	New: OB1	New: OB2	New: OB3	New: OB2 (TP)	New: OB3 (TP)	All: OB1	All: OB2
EURUSD	0			x				x
EURUSD	2			x				x
EURUSD	4	x	x			x		x
EURUSD	8			x		x	x	
EURUSD	10		x	x		x	x	x
GBPUSD	0	x	x		x	x		x
GBPUSD	2	x	x		x	x		x
GBPUSD	4		x	x	x		x	x
GBPUSD	8							x
GBPUSD	10							x

- Some consistency is found in the indicators chosen in both currencies across slippage values
- We find that the highly correlated indicators are often chosen (typically where the stop-losses are further from the spot)

- OB1 refers to sales for stop loss orders where the price is within 0.5% of the current spot
- OB2 refers to sales for stop loss orders where the price is between 0.5% & 1%
- OB3 refers to sales for stop loss orders where the price is between 0% and 1%
- TP refers to "Take Profit" orders
- Note that all indicators that were consistently not used have been removed from the above table for clarity

Conclusions

- Results show significant promise in approach
  - We demonstrate that order book based trading can significantly improve the results
- Feeding in technicals is often important - but not in isolation - rather in addition to order book information