Waverly Advisors

Backtesting and Market Research

February 25, 2014

Adam Grimes, CIO, Waverly Advisors

Outline:

- What is it and why do it?
- What skills and tools will you need?
- Specific methods
- Pitfalls and common mistakes
- Backtesting by hand

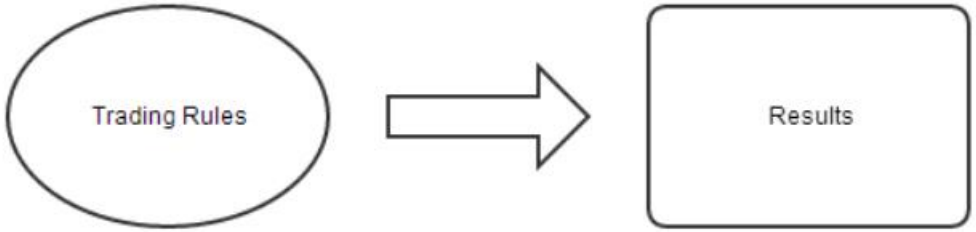
What is backtesting?

- Examining the results from following a set of trading rules using historical data.
  - May (but do not have to) consider other parameters.
  - Can (but do not have to) run “what if” scenarios?
  - Rules can be simple or complex
  - Historical data can be:
    - Long or short time periods
    - Single market or portfolio

Why backtest?

- The most common answer is to create a trading system.
  - The idea is that maybe we can find a set of trading rules that have performed well in the past and apply them in the future.
    - What assumptions are embedded in this plan?
- We can also backtest to learn.
  - I call this “market research”, to differentiate from system-creation backtest.
  - Can be:
    - General or broad
    - Market-specific or applicable to prices in general

Is It This Simple?



What Do We Need to Backtest?

- Good, clean data.
- Intelligent market selection.
- Attention to time slices.
- Robust research environment.
- Ability to collect and compare results.
- Correct statistical tools for analysis.
- The right mindset.

- None of these are easy, and all are important.

Issues with Data

- Data is rarely clean, even if you paid for it.
- Errors can range from minor to catastrophic.
- Most researchers find that the vast majority of their time (85% - 90%) is spent dealing with cleaning and wrangling datasets, and finance is no exception.
- High frequency data is especially susceptible to errors, but standard data sources sometimes include errors on daily and longer timeframes.

Examples of Data Errors

- Missing days
- Missing datapoints for some days
- Incorrect timestamps on the full series
- Incorrect timestamps on a single data point
- "Obvious" errors
  - Low < High
  - Close < 0
  - O or C > H or < L
  - Misplaced decimal point (e.g., 101.23, 103.45, 1,058.7)
- Incorrectly adjusted prices
  - Futures rolls
  - Stock prices for splits/dividends
- Mislabeled data files
- Wrong prices (e.g., 42 for 43)

Market Selection

- Some questions are easy:
  - Do you want to test an idea on futures and equities both?
- Some are trickier:
  - If you are testing currencies, do you need to test cash currency and futures both? What about exchange traded products?
- Some are really hard:
  - How many markets do you need and how much data is enough?
    - For instance, are 500 stocks better than 200?
      - What about correlation?

Market Selection

- Make sure you cover markets relevant to your trading situation and mandate.
- Reach broadly enough that you can examine the "edges" of your question.
  - Do all markets trade the same, as technical analysis teaches?
- Make an effort to include different market conditions and volatility regimes.
  - Bull, bear, and flat markets.
  - Quiet periods and highly volatile periods.
  - Periods with and without exogenous stresses.

Research Environment

- There are many pieces of software that claim to give tools for backtesting.
- Many blogs, of differing degrees of sophistication, giving ideas and guidelines.
  - When you are starting out, do not assume that everything you read is true.
  - You need to develop knowledge and skills to discriminate.

Clarify Your Objectives

- "Back of the envelope" testing.
- Building complete trading systems
- Creating HFT algos
- Interfacing directly with execution software
- Collecting market data (e.g., tick level)
- Developing your trading skills/intuition

- These are all very different and will require different tools.
- The right tool for one job can be useless for another.

Options and Possibilities

No	Adequate	Good	Probably overkill
Retail-level platforms	Excel (Pencil and paper?)	Excel + VBA Python R Matlab, etc Stata, SAS, SPSS Java	C++

Special Mention: Excel

- Excel is a very useful tool, and many people get their first steps in quantitative analysis in Excel.
- There are good advantages:
  - Easy to visualize data.
  - Easy to manipulate data.
- But also disadvantages.
  - Slow
  - Lacks standard graphing capabilities (histograms)
  - Can be very cumbersome when projects get complicated.
- Is much more powerful when combined with VBA.
- My advice: learn it. Use it, but don't depend on it: Excel is probably not your final solution.

Programming Skills

- You are probably going to need to develop some programming skills.
- Can range from relatively simple, application specific (i.e., being able to operate a piece of software), to learning a complete language.
- If this scares you, don't be intimidated. It's a skill, like any other, than takes time to develop, but anyone can learn to do it.
  - There are some very good online resources.

Analyzing Results

- You have to be able to understand the statistics.
- For example:
  - What measure of central tendency is relevant?
  - How do you need to factor dispersion in?
  - What significance test (if any) is relevant?
  - How strong is the effect?
  - How stable is it?
- The simplest output might be three numbers: a mean, standard deviation, and sample size.
- There is no way around needing these skills, but, again, there are many free resources.

Mindset

- It can be constructive to take an objective, scientific approach to market problems.
- The scientific method and research protocols from other disciplines might be good models.
- There is also value in market experience (street smarts) and common sense.
- These are gained through experience and exposure to market data.
- Good research methodologies incorporate an element of art.
- Correct research methodologies can aid in the development of intuition and market sense.
- These are not opposed!

Discretionary vs. Quantitative

- This is not a division. Rather, it is a spectrum.
- There are purely quantitative traders.
  - Algorithmic traders.
  - System traders.
- Are there purely discretionary traders?
  - A successful, purely discretionary trader will, at the very least, draw on his own experience.
    - This is a form of backtesting and learning.
- Many traders blend the two.
  - Need rules.
  - Discipline is essential.

Specific Tools

- Simple pattern testing, using event study methodology.
  - *Event Studies in Economics and Finance*, MacKinlay, 1997

- Portfolio-level testing platform
  - Produce portfolio metrics
  - Should include some ability to optimize

Optimization: Not a Dirty Word

- Optimization is a technique that takes many different inputs (e.g., length of moving average) and finds which would have produced the best results.
  - (It is not “curve fitting”. That term is mis-used by technical analysis!)
- Over-optimization is bad. With enough parameters, virtually any trading system can be made to create impressive results on historical data.
  - “With four parameters I can fit an elephant and with five I can make him wiggle his trunk.” - John von Neumann
- There is a good word for the right optimization: **Learning.**

Optimization: Guarding Against Mistakes

- Backtesting results should always be suspect, but optimized backtest results especially so.
- (Errors can arise even when there is no intent to deceive.)
- Simple guidelines can help:
  - Do not optimize heavily.
  - Be suspicious of results that show a single impressive result in many poor tests.
    - Look for “areas” or “zones” of good results.
    - Maybe select a sub-optimal parameter set.
  - **Use out of sample data correctly.**

Out of Sample

- Out of sample data is data that is held back and not touched at all in the research process.
- The workflow is that you create your system, test it, refine it, and, as a last step, test it on out of sample data.
- You only get one shot!
  - This is **the** critical aspect of out of sample testing, and one that is often ignored.
- Once the out of sample has been used, even once, it is now in sample and **cannot** be used again.
- A question: if you are trading actively and experience market data, can it be used as out of sample?

"Sneaky" Optimization

- Imagine a process like this:
  - What happens if we put moving average X on a market? (2,000 trades)
  - What if we add indicator A? indicator B? indicator C? Let's keep indicator B and C. (500 trades)
  - What if we only traded this on Mondays? Tuesdays? The results are so much better if we eliminated Wednesdays, so let's do that. (100 trades)
  - What if we add a trend indicator? (50 really good trades…)

- This is a process of optimization, and can easily become over-optimization.
- Be careful of a process that starts with many events and filters down to "only the best trades".

Be Careful of Ratio Metrics

- When comparing different results, what should you compare?
  - How much money you would've made?
  - How many trades?
  - How variable those results were?
  - Winning percentage?
- A typical quant module might simply compare Sharpe ratios of optimization runs, and pick the highest Sharpe.

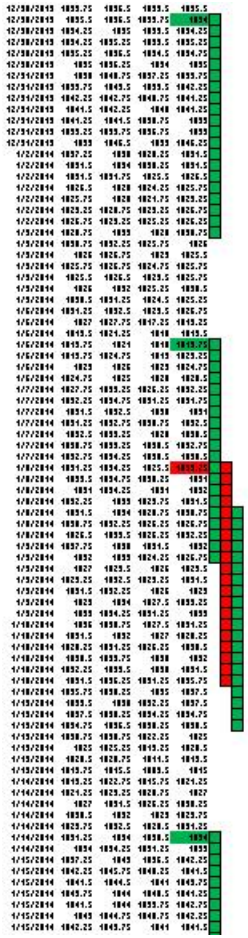
Sharpe Ratio

- The concept is average return divided by standard deviation of those returns.
- Possible issues:
  - Does “average return” describe everything that needs to be considered?
  - Is standard deviation a good measure of risk?
  - Are there other factors to consider?
- In general, ratio metrics are questionable because they are very sensitive to changes in the denominator.
- There are other ratios and other metrics to consider.

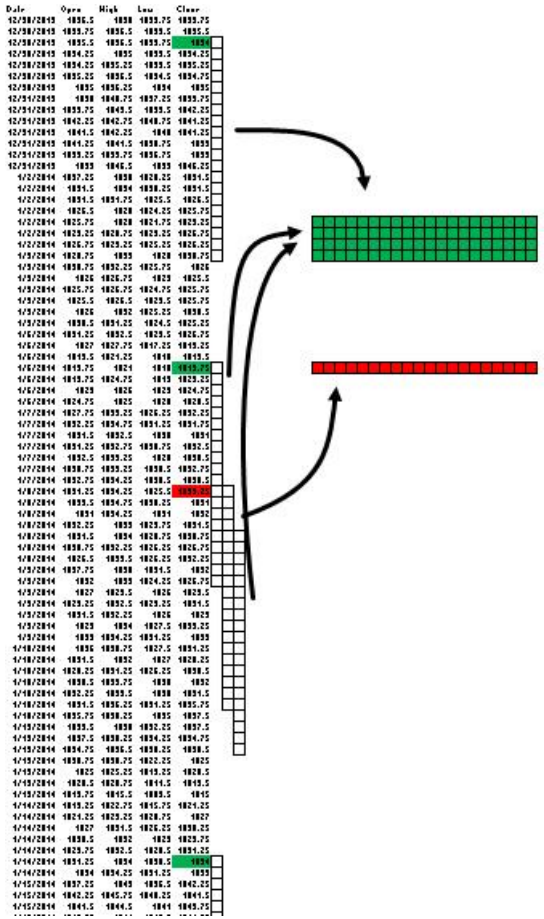
Event Study

- Define a precise condition that will result in a trading signal.
  - Symmetrical for buy and sell.
- Work through each bar of the universe looking for the condition.
- Record returns for each bar following the condition.
- Combine all N+1, N+2, … returns to get a composite for each bar following the signal.
- Create excess return measure (signal – baseline) for each bar.

Event Study



Event Study

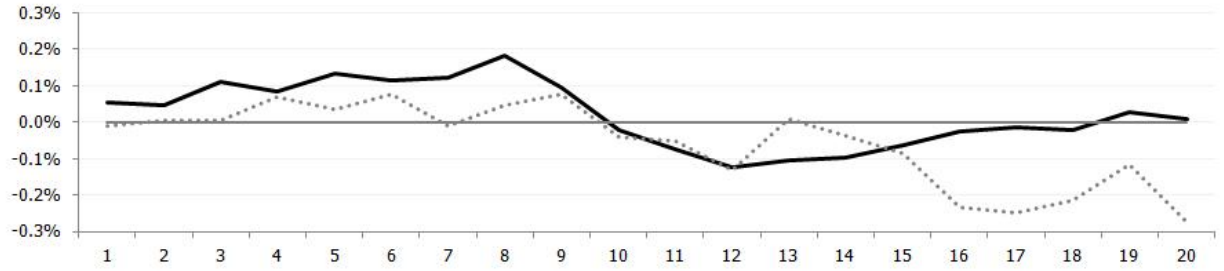


Event Study Output

		Raw			Excess				Baseline				
Day	N	Up|Down	Mean	Median	Mean	StDev	Median	IQR	Up|Down	Mean	StDev	Median	*p-val*
1	3,347	50.7%|48.3%	0.11%	0.04%	**0.05%**	2.69%	(0.01%)	2.10%	50.0%|48.5%	0.06%	5.31%	0.01%	*0.00*
2	3,347	51.5%|47.4%	0.13%	0.09%	**0.05%**	3.89%	0.00%	3.08%	51.1%|47.9%	0.08%	19.40%	0.08%	*0.00*
3	3,347	51.9%|47.1%	0.23%	0.13%	**0.11%**	4.63%	0.01%	3.67%	51.8%|47.4%	0.12%	30.17%	0.14%	*0.00*
4	3,347	52.5%|46.9%	0.24%	0.23%	**0.08%**	5.26%	0.07%	4.43%	52.2%|47.0%	0.16%	49.49%	0.20%	*0.00*
5	3,347	52.2%|47.0%	0.30%	0.21%	**0.13%**	6.71%	0.04%	4.88%	52.6%|46.7%	0.17%	68.40%	0.26%	*0.00*
6	3,347	53.4%|46.3%	0.38%	0.34%	**0.11%**	7.20%	0.08%	5.35%	53.0%|46.3%	0.26%	69.27%	0.33%	*0.00*
7	3,347	53.7%|45.7%	0.47%	0.34%	**0.12%**	8.45%	(0.01%)	5.77%	53.3%|46.0%	0.35%	68.21%	0.39%	*0.00*
8	3,347	53.5%|46.0%	0.57%	0.43%	**0.18%**	9.08%	0.04%	6.36%	53.7%|45.7%	0.39%	67.27%	0.46%	*0.00*
9	3,347	54.6%|45.2%	0.59%	0.57%	**0.10%**	8.57%	0.08%	6.69%	54.0%|45.3%	0.49%	62.93%	0.53%	*0.00*
10	3,347	53.8%|45.7%	0.57%	0.55%	**(0.02%)**	8.86%	(0.04%)	7.06%	54.2%|45.1%	0.59%	56.12%	0.58%	*0.00*
11	3,347	54.6%|45.1%	0.62%	0.64%	**(0.07%)**	9.16%	(0.05%)	7.41%	54.5%|44.8%	0.70%	52.75%	0.65%	*0.00*
12	3,347	54.8%|45.0%	0.67%	0.66%	**(0.12%)**	9.26%	(0.13%)	7.71%	54.7%|44.6%	0.79%	54.05%	0.72%	*0.00*
13	3,347	55.5%|44.1%	0.77%	0.89%	**(0.11%)**	9.56%	0.01%	7.90%	54.9%|44.5%	0.88%	60.92%	0.78%	*0.00*
14	3,347	55.5%|44.2%	0.81%	0.87%	**(0.10%)**	10.10%	(0.04%)	8.20%	55.1%|44.2%	0.91%	63.84%	0.84%	*0.00*
15	3,347	55.5%|44.4%	0.88%	0.86%	**(0.06%)**	10.64%	(0.09%)	8.83%	55.3%|44.0%	0.94%	66.25%	0.91%	*0.00*
16	3,347	54.6%|45.1%	0.95%	0.74%	**(0.02%)**	11.92%	(0.23%)	9.01%	55.4%|43.8%	0.98%	77.18%	0.97%	*0.00*
17	3,347	54.5%|45.1%	1.00%	0.77%	**(0.01%)**	12.24%	(0.25%)	9.21%	55.7%|43.6%	1.01%	80.28%	1.04%	*0.00*
18	3,347	55.4%|44.4%	1.07%	0.88%	**(0.02%)**	12.65%	(0.21%)	9.43%	55.8%|43.4%	1.10%	66.84%	1.10%	*0.00*
19	3,347	55.3%|44.4%	1.16%	1.02%	**0.03%**	13.07%	(0.12%)	9.83%	55.9%|43.3%	1.13%	70.37%	1.15%	*0.00*
20	3,347	54.7%|45.0%	1.20%	0.91%	**0.01%**	14.19%	(0.28%)	10.24%	56.0%|43.2%	1.19%	68.67%	1.20%	*0.00*

Baseline: Up: 50.0%, Down: 48.5%, Unch: 1.5%

Mean (solid) +/- 1σ and Median (dashed) Excess Returns by Bar Following Event



Example Workflow for Research

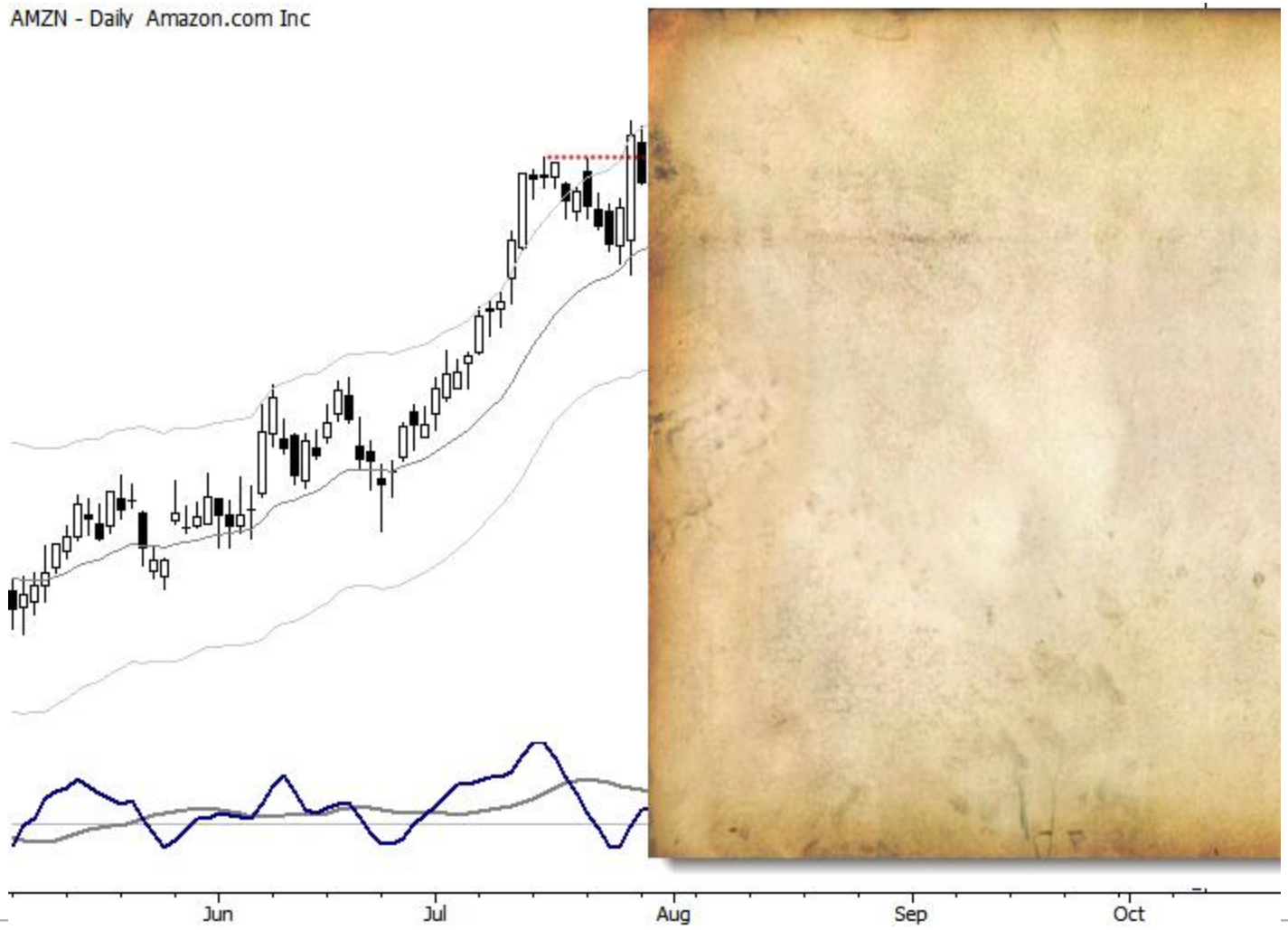
1. Formulate an idea about market action. Be precise.
2. Come up with a plan for testing the idea.
3. Gather correct data and tools for doing the analysis.
4. Analyze the data.
5. Analyze the results.

- Refine the question
- Optimize
- Use out of sample

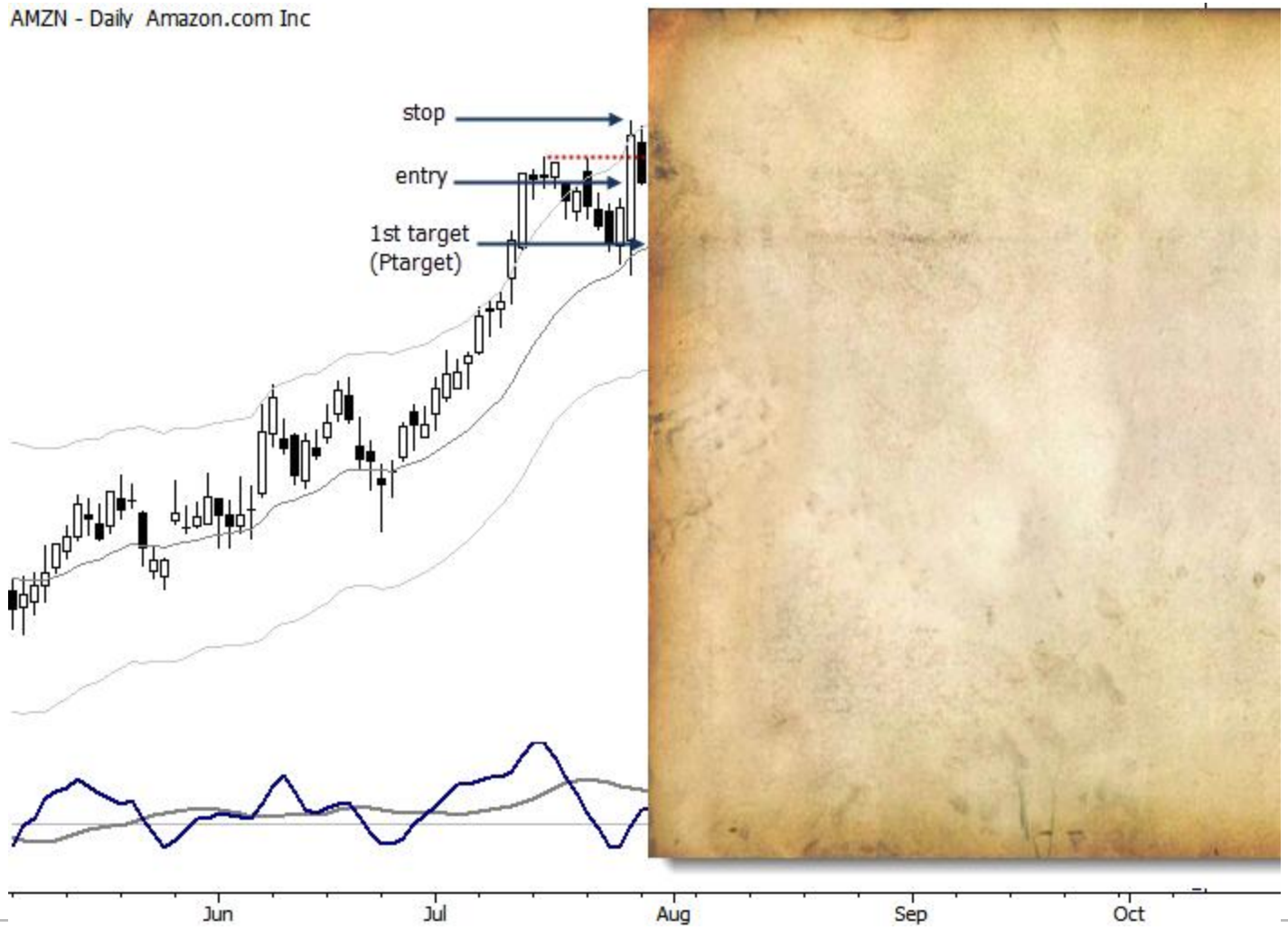
Manual Backtesting

- Automated backtesting has the advantage of breadth and precision.
  - Can look at many markets and hundreds of thousands of events.
  - Can easily examine different time periods and markets for stability.
- Manual backtesting, by contrast, is slow and may appear to be cumbersome.
  - But close contact with market data helps build intuition and market sense.
- Bar by bar testing is a good tool, but there are some pitfalls to consider.

Bar by bar: Entry



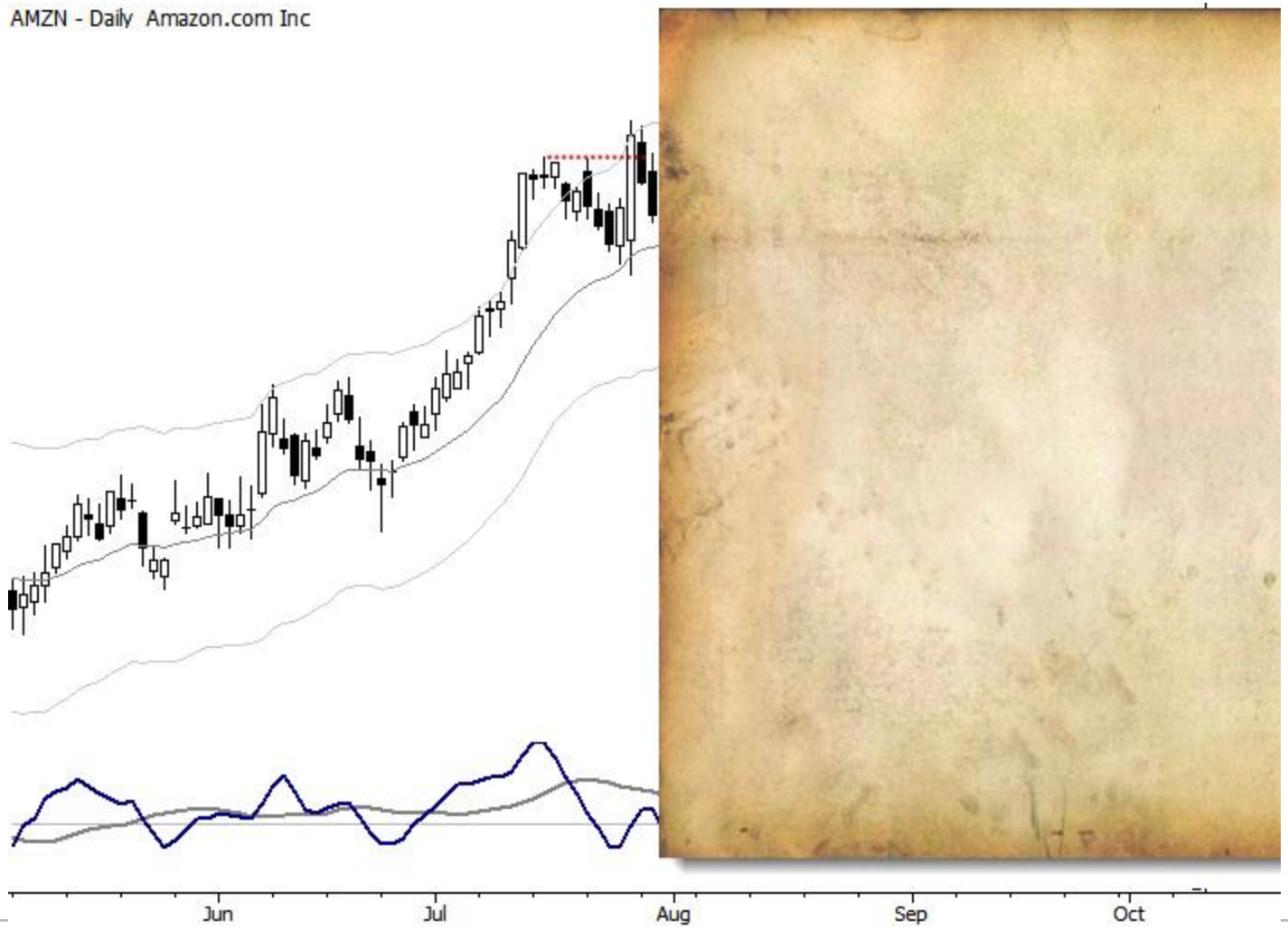
Three Prices



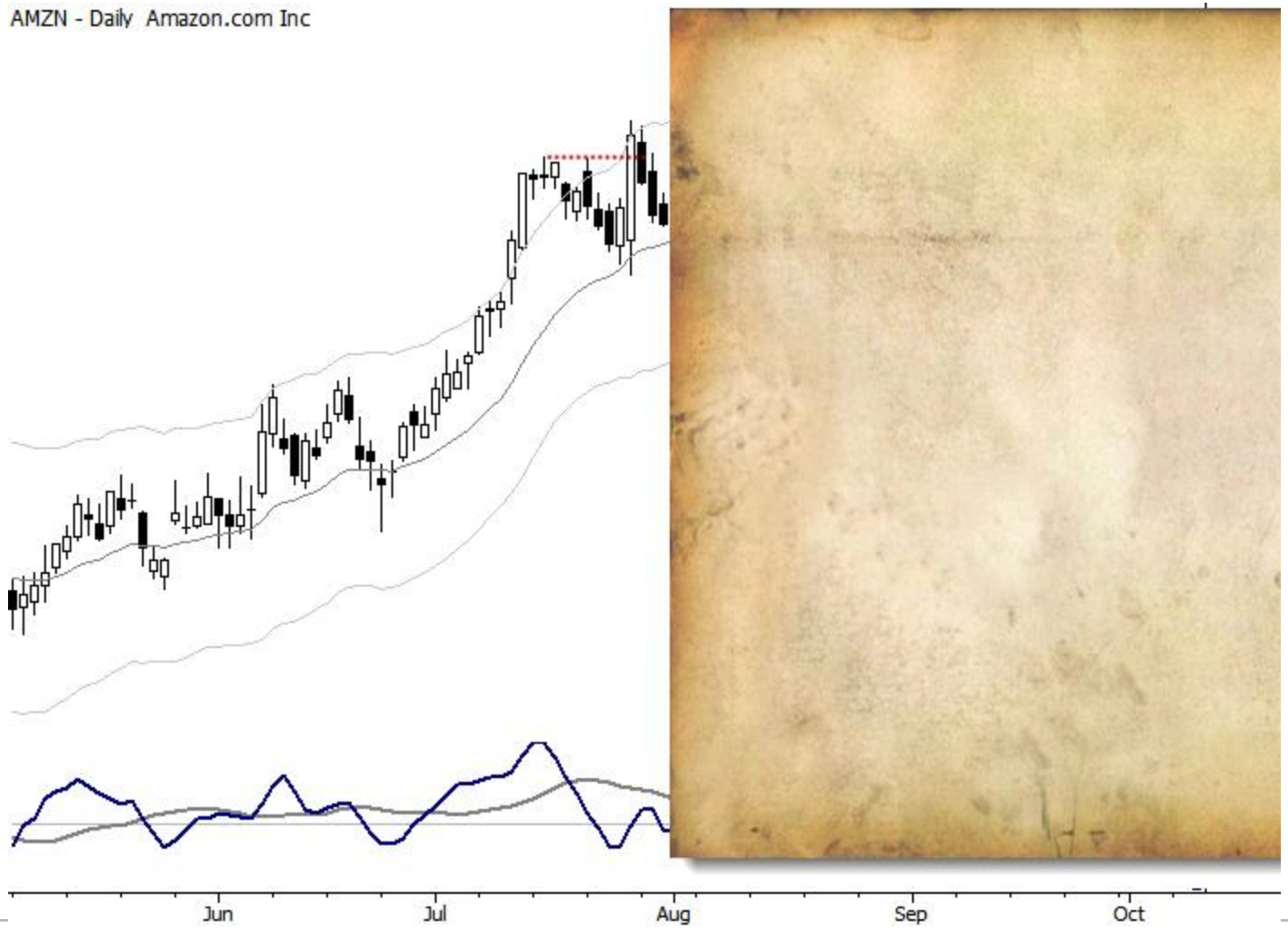
P&L Sheet Entry

	A	B	C	D	E	F	G	H	I	J	K	M	N	O
1		Date	Market	Type	Mult	L/S	#	PriceIn	PriceOut	InitStop	P&L	%R	Ptarget	Target?
2		7/29/2013	AMZN	FT	1	-1	1	306.10		313.62	[illegible]	[illegible]	298.58	
3														

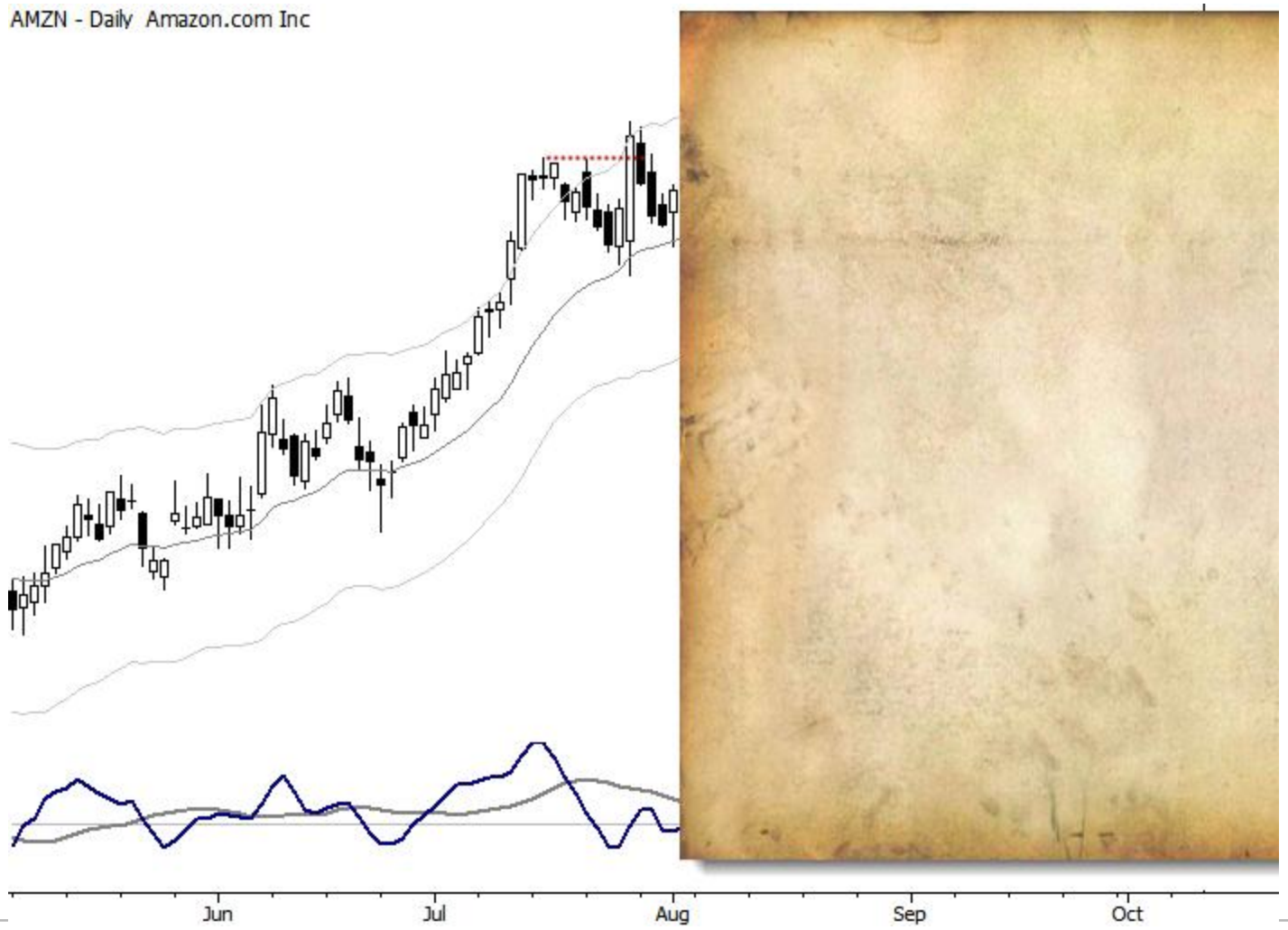
Bar by Bar…



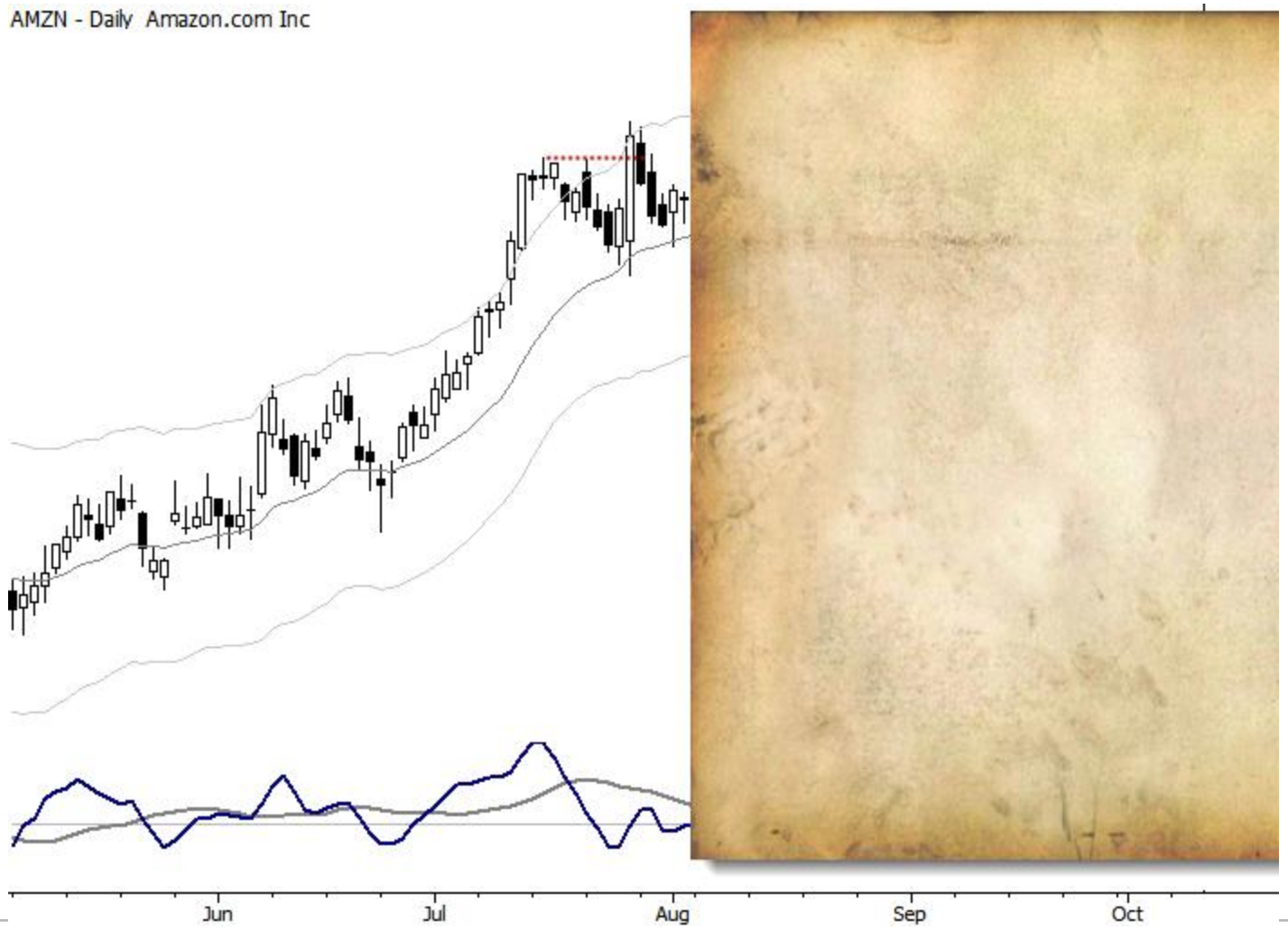
Bar by Bar



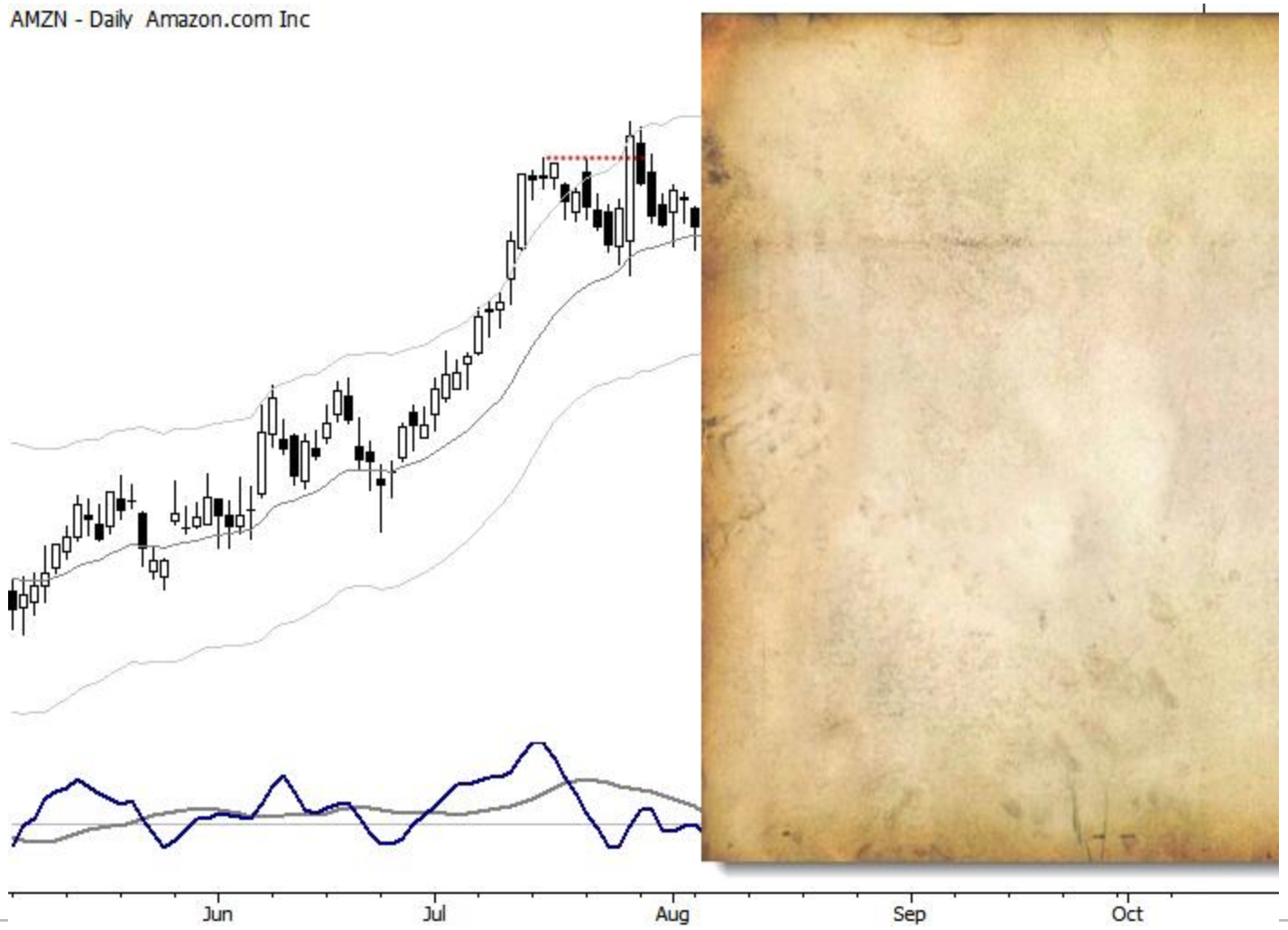
Bar by Bar



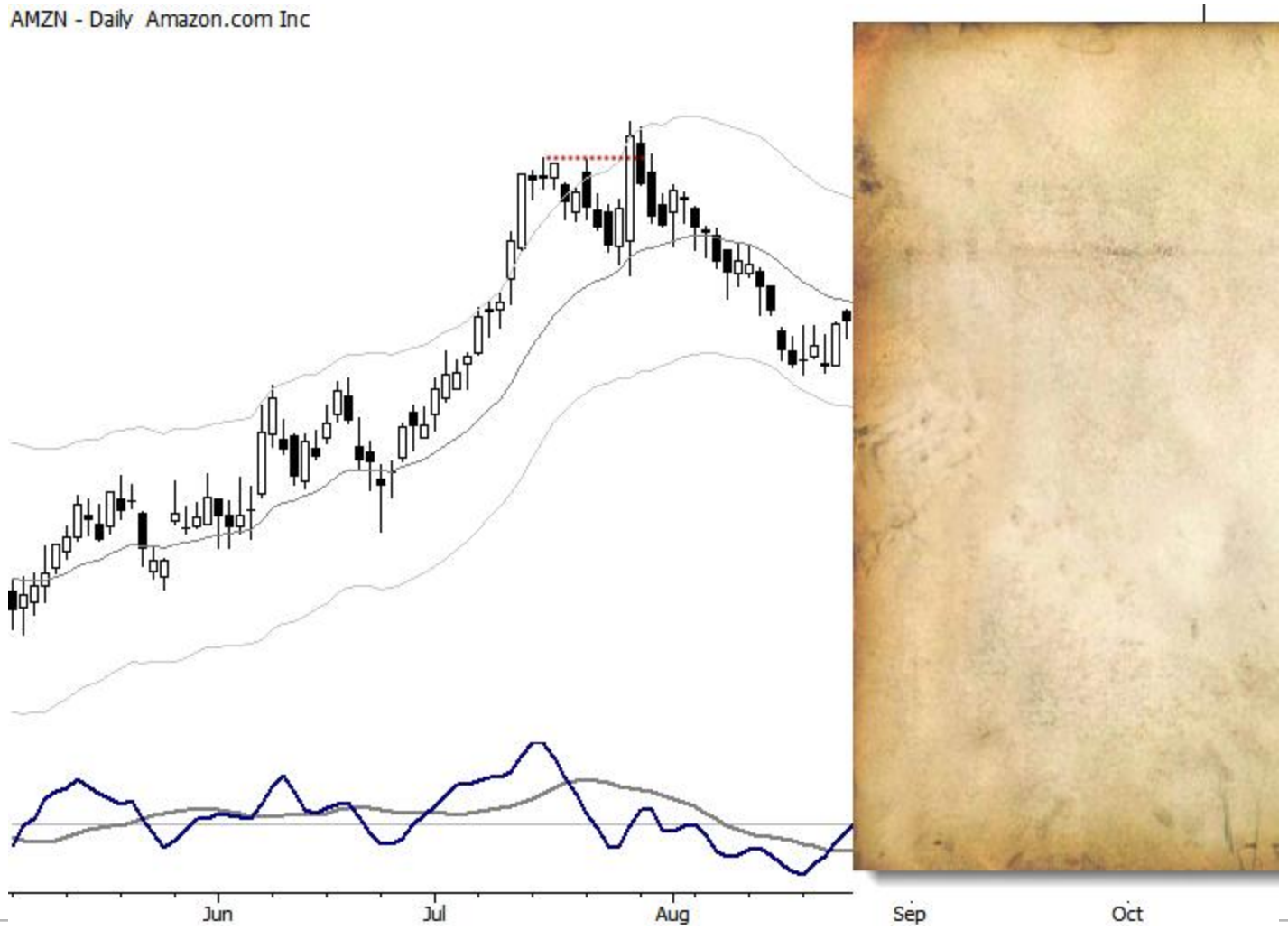
Bar by Bar



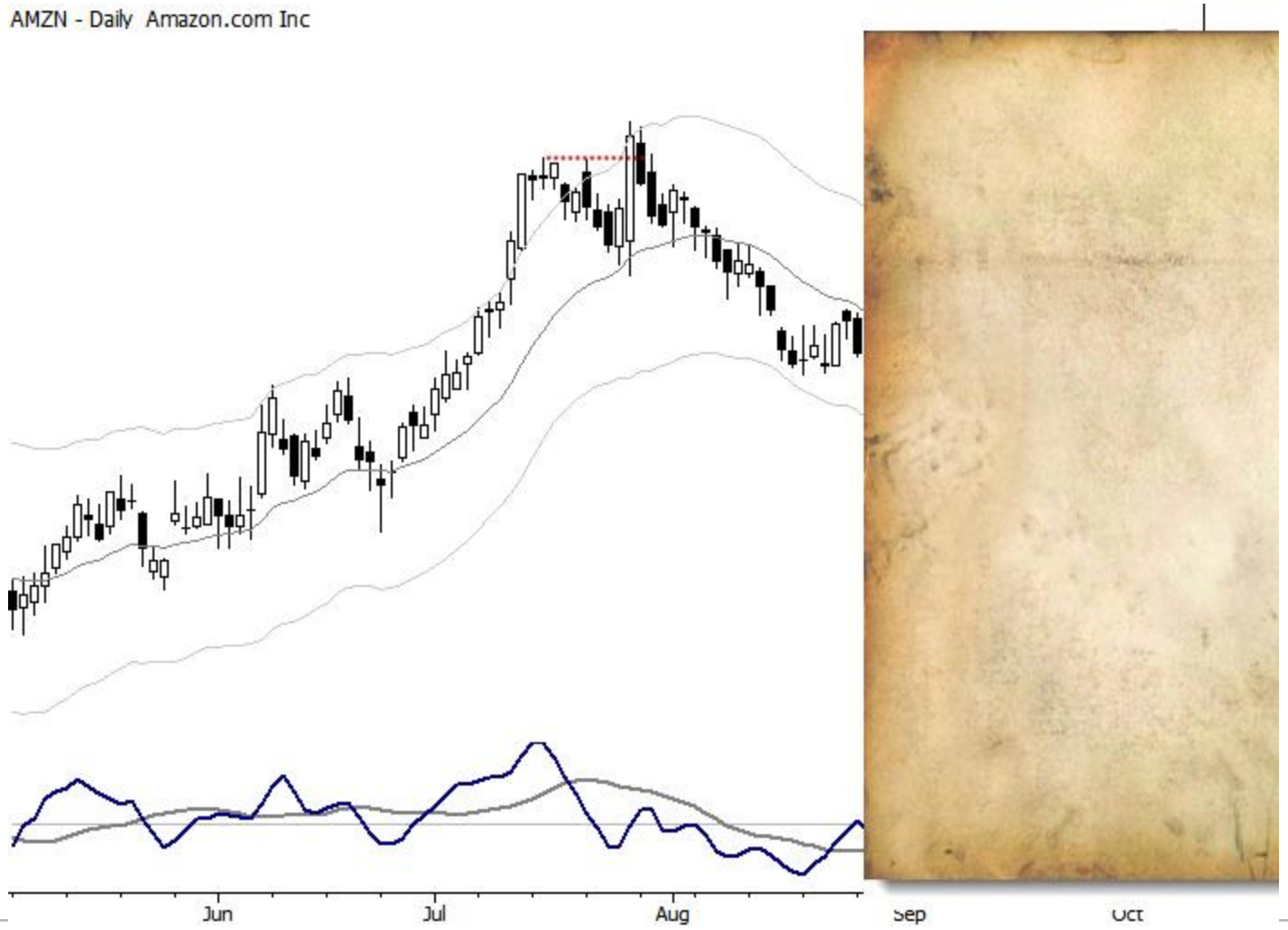
First Target Hit



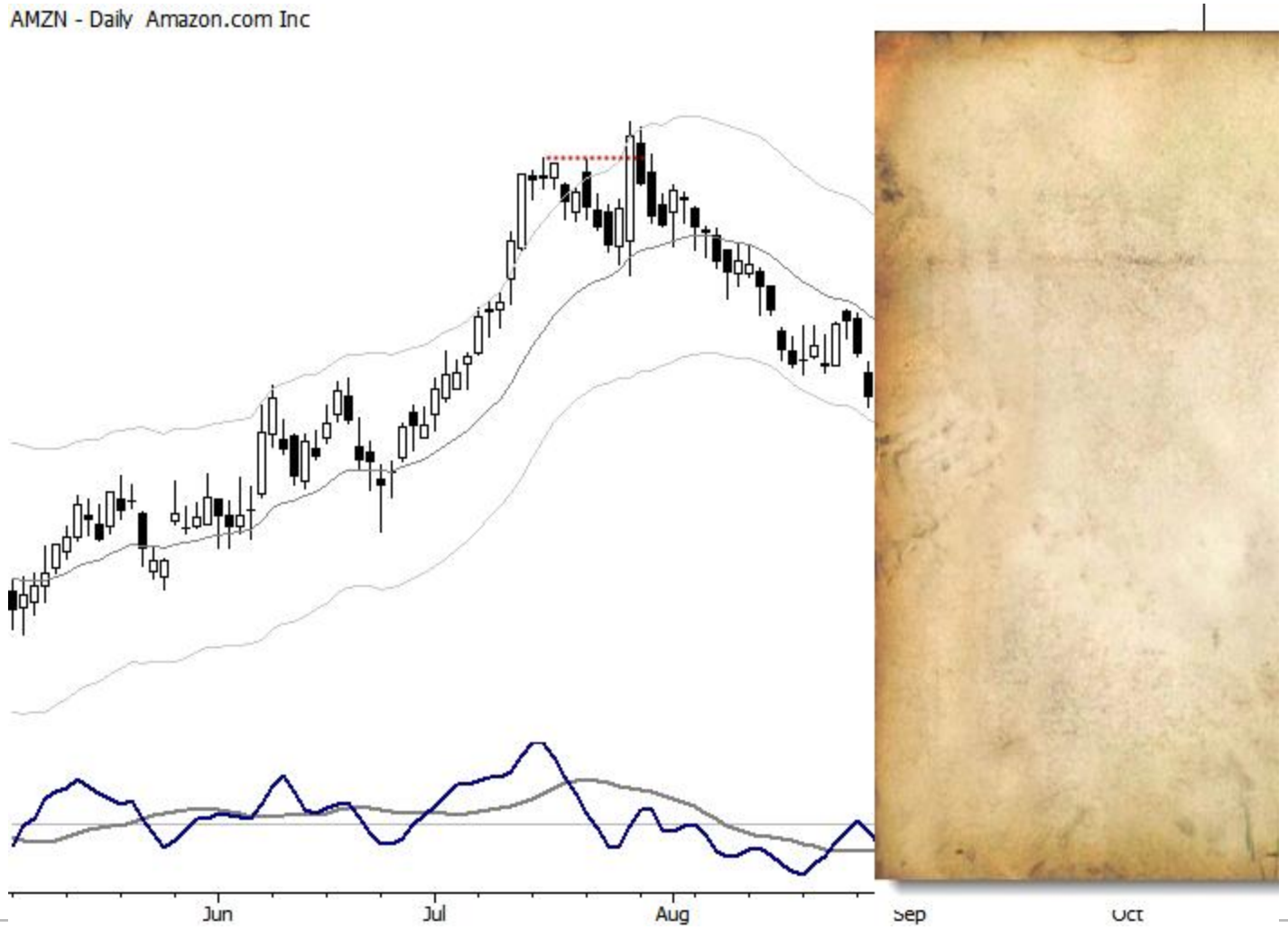
What is the Plan?



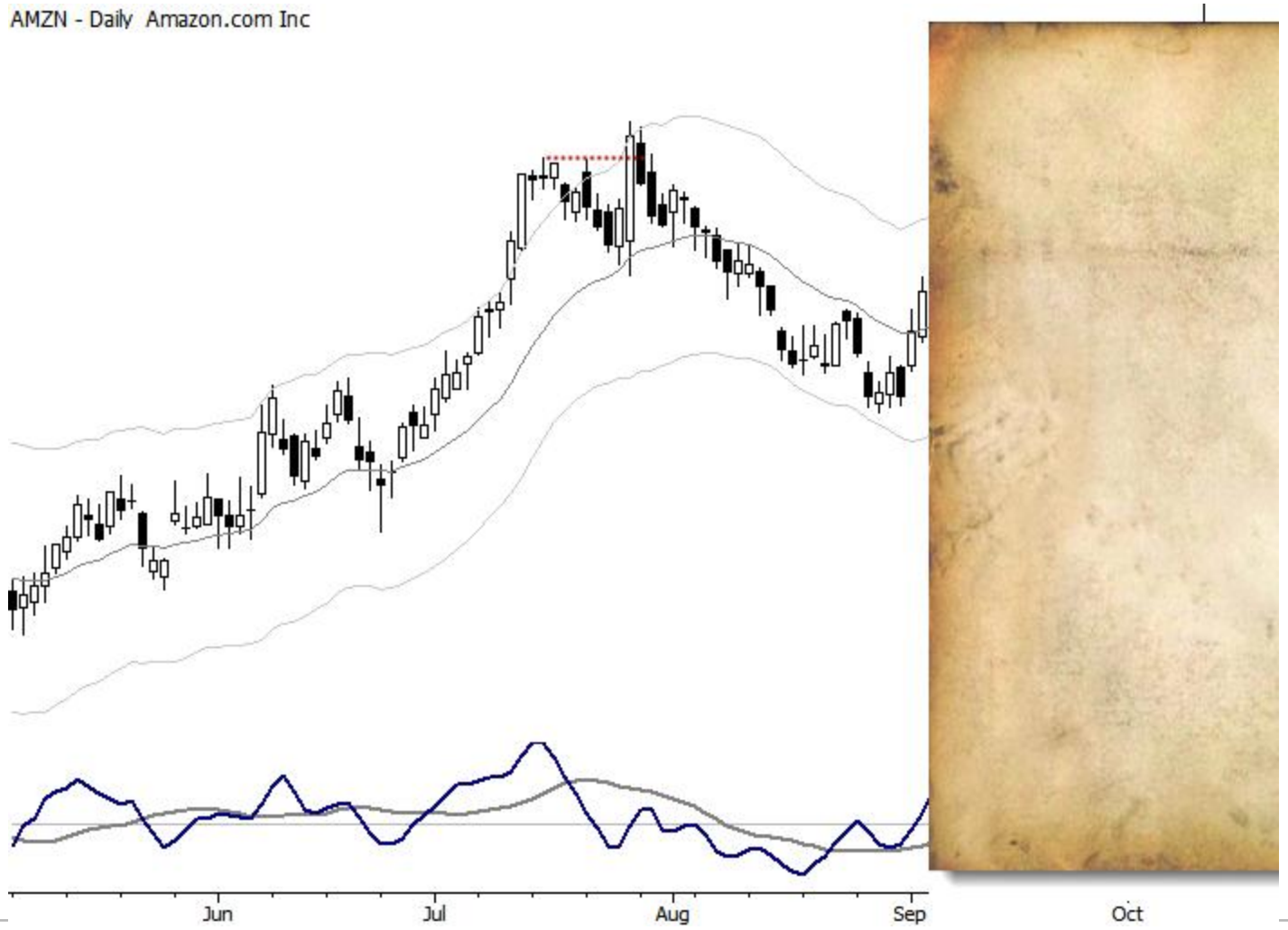
Bar by Bar



Bar by Bar



Continue Until You Exit

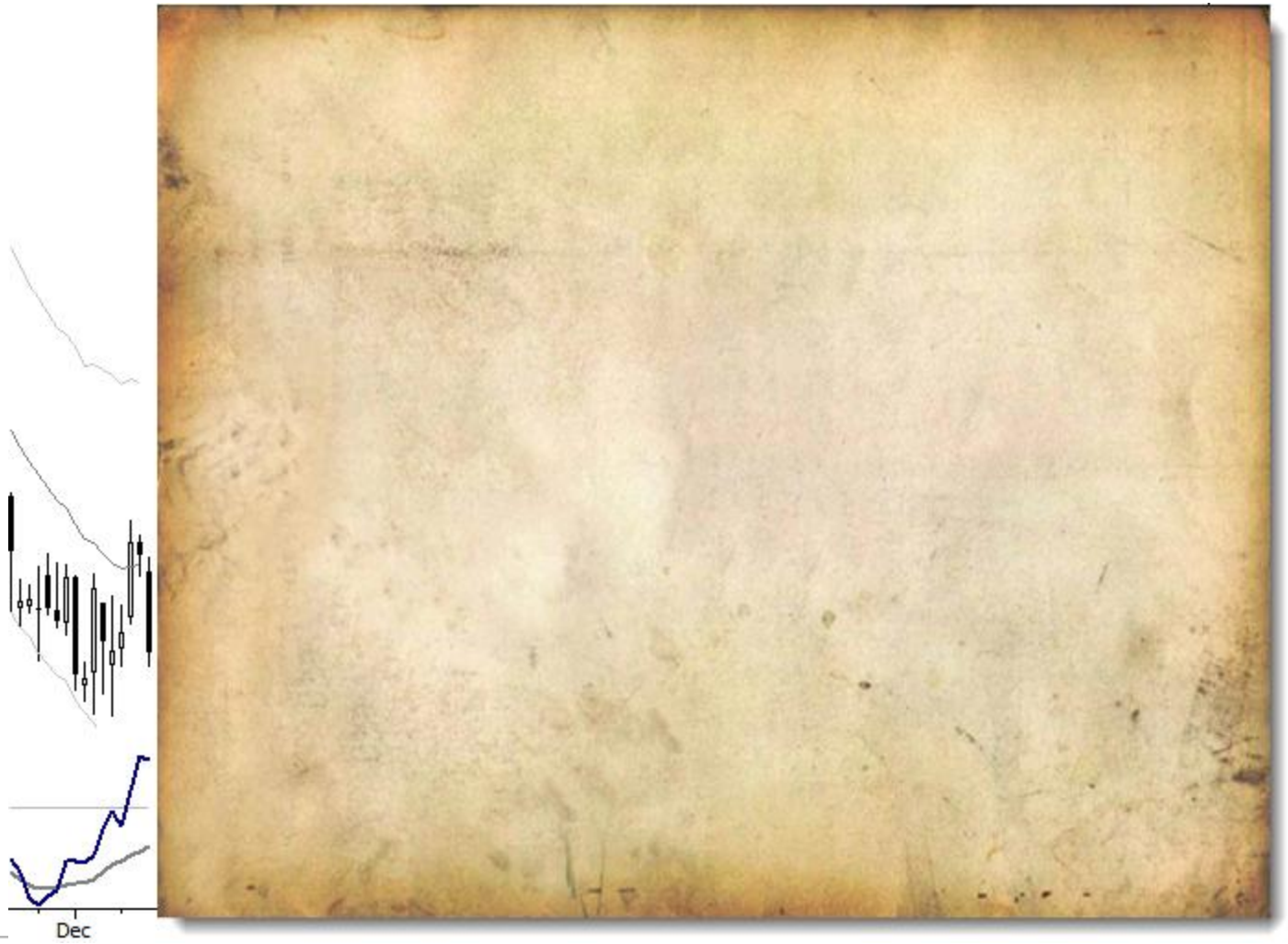


Note Exit Price in Sheet

	A	B	C	D	E	F	G	H	I	J	K	M	N	O
1		Date	Market	Type	Mult	L/S	#	PriceIn	PriceOut	InitStop	P&L	%R	Ptarget	Target?
2		7/29/2013	AMZN	FT	1	-1	1	306.10	293.64	313.62	12.46	1.7x	298.58	1

But there is a problem…

Bullish or Bearish Bias?



How About This One?

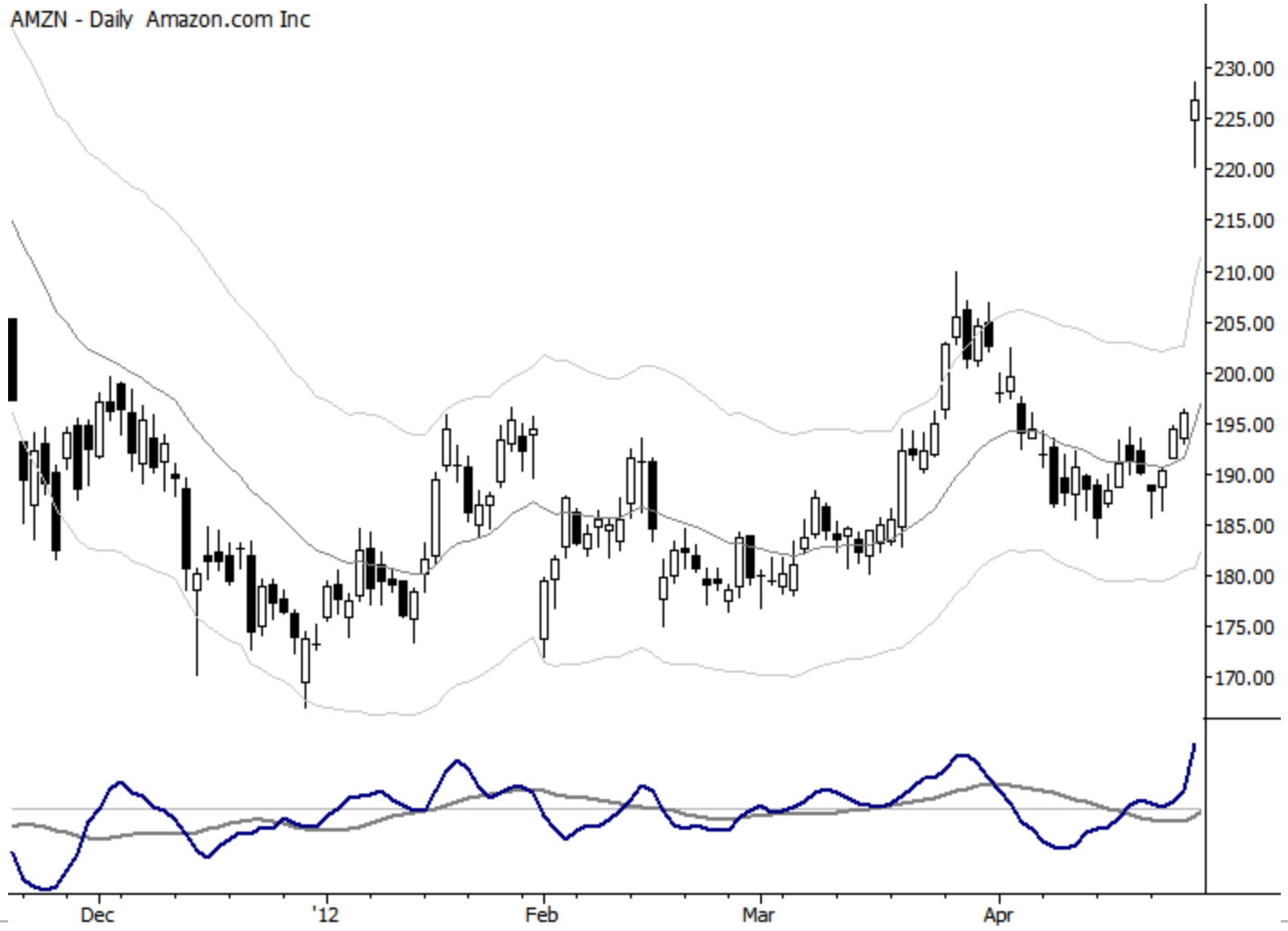
The Cardinal Rule

- In these tests, you must not access any information that you would not have had in real time.
- Any “leakage” will severely compromise your results.
  - Data snooping.
- Consider your decisions.
  - If you modify a decision even once, you have invalidated the whole process.
- Is there any way you are getting information from the future?

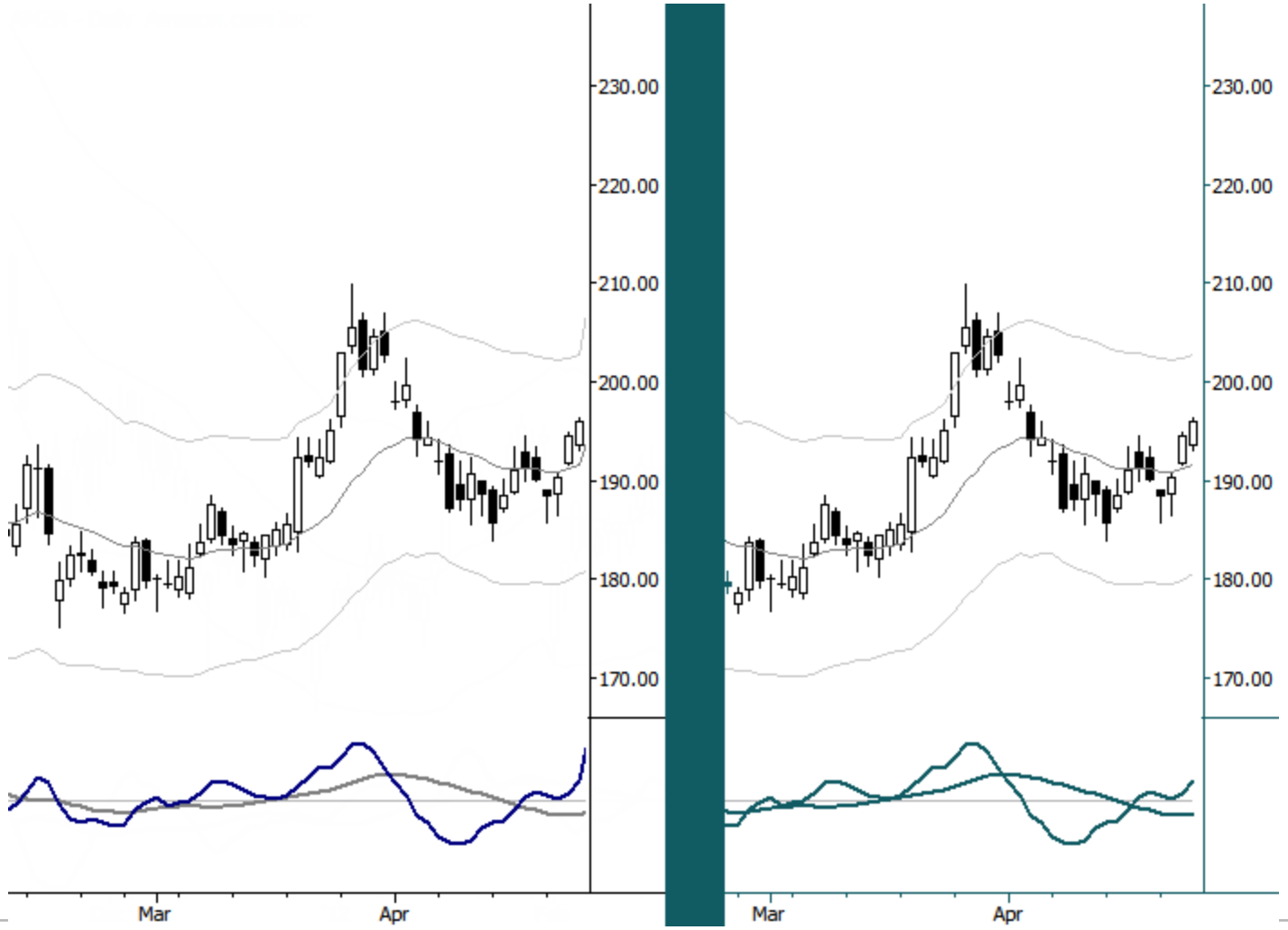
A Better Way

- Go bar by bar in your software.
- Even then, be careful.
  - Do you see hints of the next bar?
  - Can you get hints from your indicator lines?
- Do not underestimate the impact of leakage.

Consider this Gap



Would You Make a Different Decision?

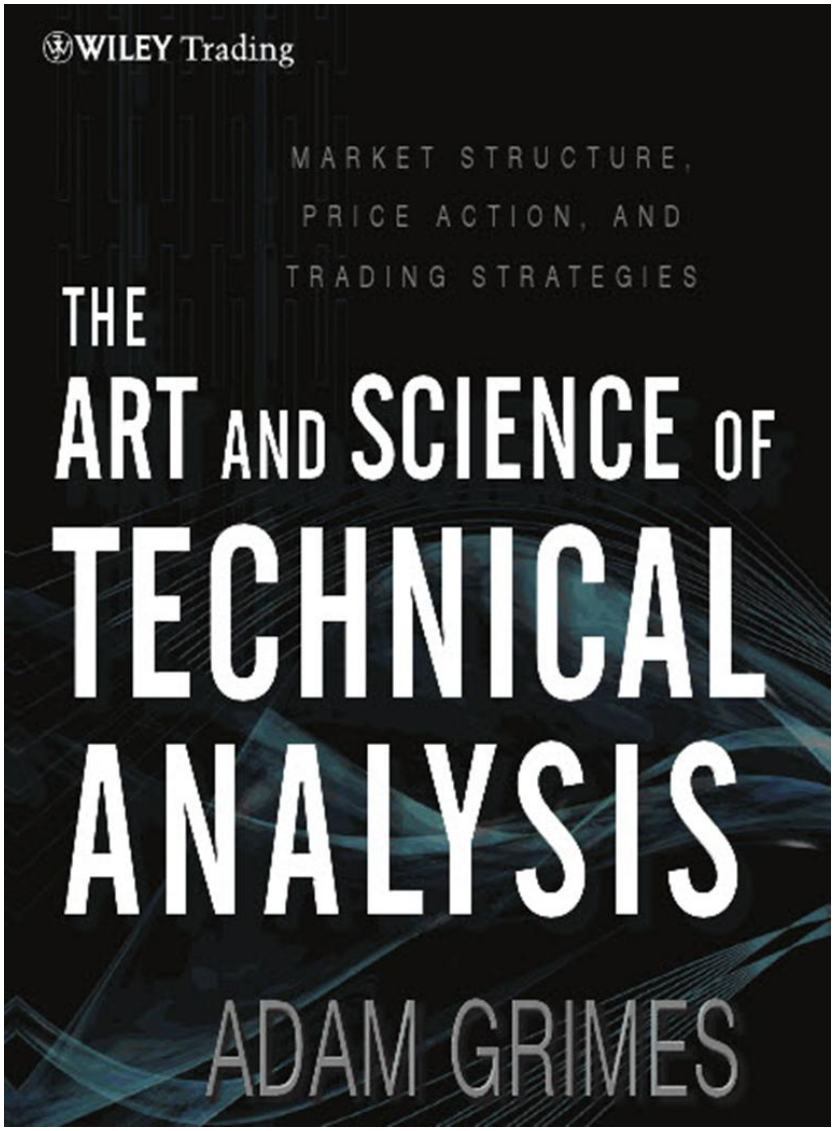


Questions to Ask

- What is your exact entry price?
- Where is your stop?
- Are you using a profit target?
- What decisions would you have made bar by bar?
- Where is your final exit?
  - If multiples, what is average price?

Consider Manual Backtesting

- Done well, this can replicate some of the experience of decision making under uncertainty.
- Done well, this process can cross the line between backtesting and papertrading.
- Keep good records.
- Be patient.
- Be disciplined.



Waverly Advisors' Process

- Our methodology is blended quantitative / discretionary.
- Quantitative work was both high-level and system-specific.
  - For breadth, went all the way back to commodity prices in the Middle Ages. It work.
- An interesting twist: our research was done to verify an existing, successful methodology, not to create one.
- The discretionary component is the input of a trader (me!) who has spent nearly every day of the past 20 years watching markets and prices, actively managing risk for much of that time.

Waverly Advisors' Research

- Specific systems, broad tendencies, and actionable ideas in major liquid markets.
  - Futures
  - Currencies
  - Stocks (indexes and individual names)
- Both trend-following and counter-trend components.
- Applicable to traders working on all timeframes.
  - Daytraders—swing traders—investors

Waverly Advisors, LLC: Research Products

Tactical Playbook – Available on Interactive Brokers

- Written for the active trader on the daily/weekly timeframes
- Exact trade recommendations
  - Hybrid systematic-discretionary methodology
- In-depth technical "drill down" into a set of markets.
- Bigger-picture overview of all liquid asset classes.

Tactical Portfolio Outlook – Available on Interactive Brokers

- Written for the longer-term manager
  - Addresses both the allocator and the longer-term active trader.
- Emphasis on executing with ETFs in a long-only and long-short environment
- Focus on Equities, Equity Sectors, and other asset classes
- Macro perspective on risk factors and major economic events.

Options Market Outlook – Contact Waverly Directly

- Proprietary, quantitative analysis of options market
- Incorporates both volatility and directional analysis
- Macro risk factors and cross-asset perspective
- Actionable trade ideas

Contact:

Waverly Advisors
228 Cedar Street
Corning, NY 14830

(607) 684-5300

www.waverlyadvisors.com
info@waverlyadvisors.com

Adam Grimes
CIO, Quantitative Analysis, Risk Management
grimes@waverlyadvisors.com

Andrew Barber
CEO, Macroeconomic Analysis
barber@waverlyadvisors.com

Chris Noye
Managing Director, Head of Sales
noye@waverlyadvisors.com