Combining Managerial Insights with Predictive Models when Setting Optimal Prices

Alan L. Montgomery
Professor of Marketing, Carnegie Mellon University
Visiting Professor, HKU

E-mail: alanmontgomery@cmu.edu

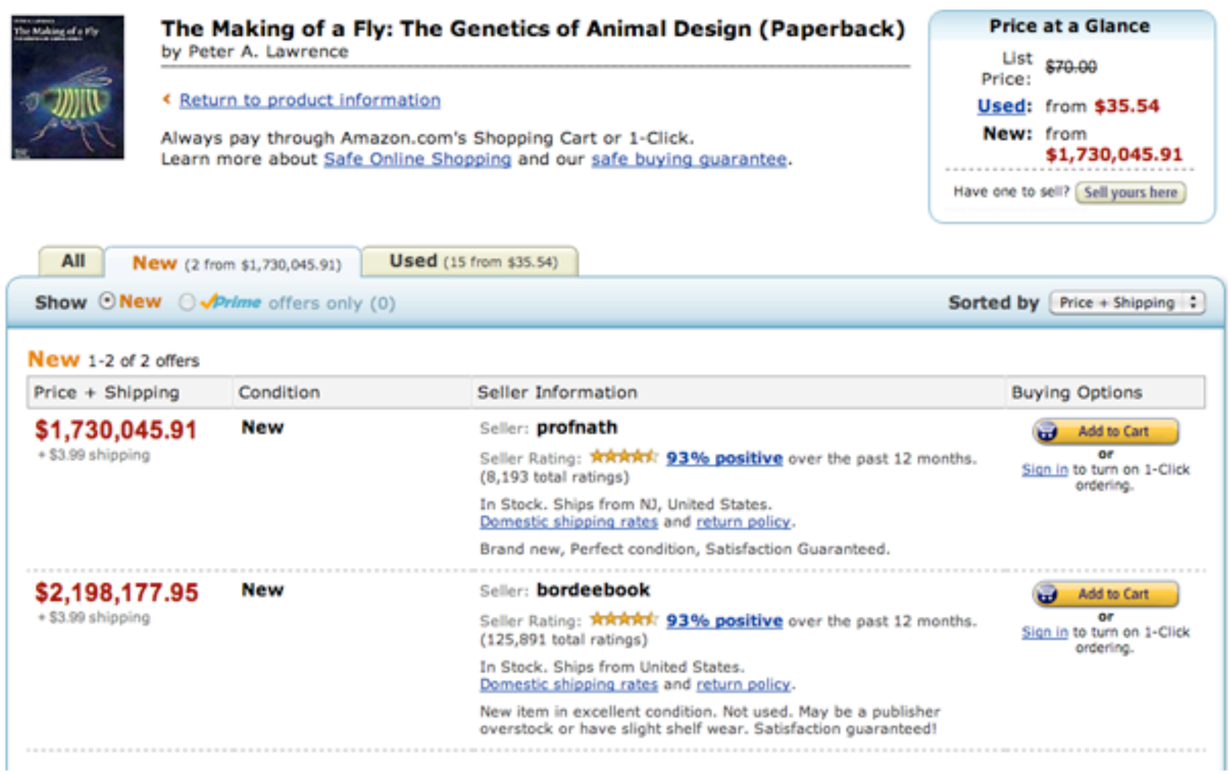
Research Goal

A popular feature of both optimal pricing solutions in both industry and academics is to implement business rules which impose constraints on the optimal pricing solution.

These constraints reflect prior information on the part of the user about the pricing solution. However, current approaches treat this information in an ad-hoc manner.

This research shows how business rules can be formulated as prior information to yield *better* pricing decisions.

Example
Why did Amazon charge $2m for textbook?



Example

What is the optimal price for "Asparagus Water"?

Broader Research Question

How to combine Data Science/ML with Human/Expert/ Managerial Intuition/Judgment?

- Only use the data (objective but inefficient)
- Only use managerial insight (subjective but explainable)
- Use data unless conflicts with managerial insight
- Use managerial insight unless it conflicts with data
- Combine both together (weight each appropriately)

Outline

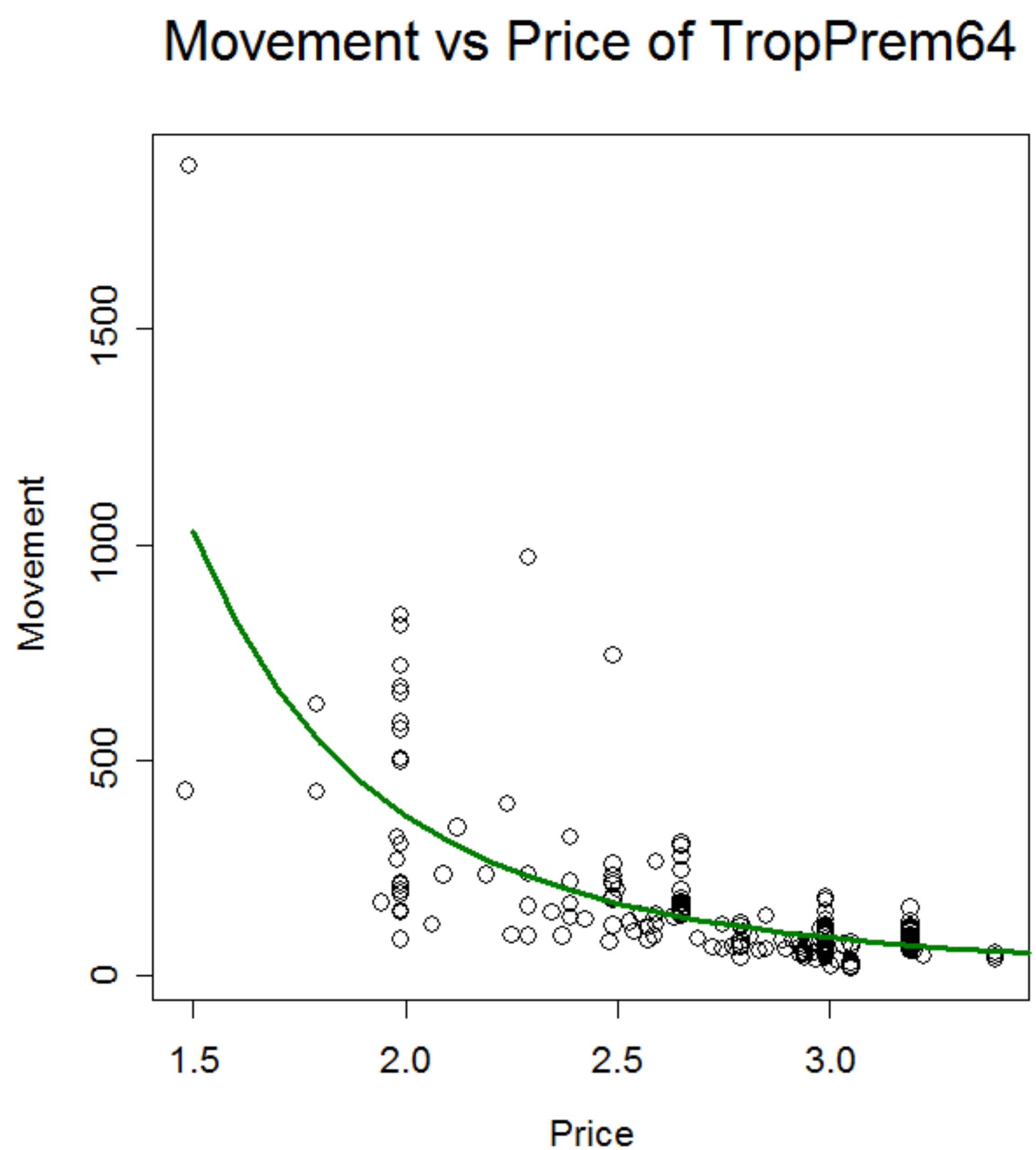
Current practice of price optimization

Business rules as prior information

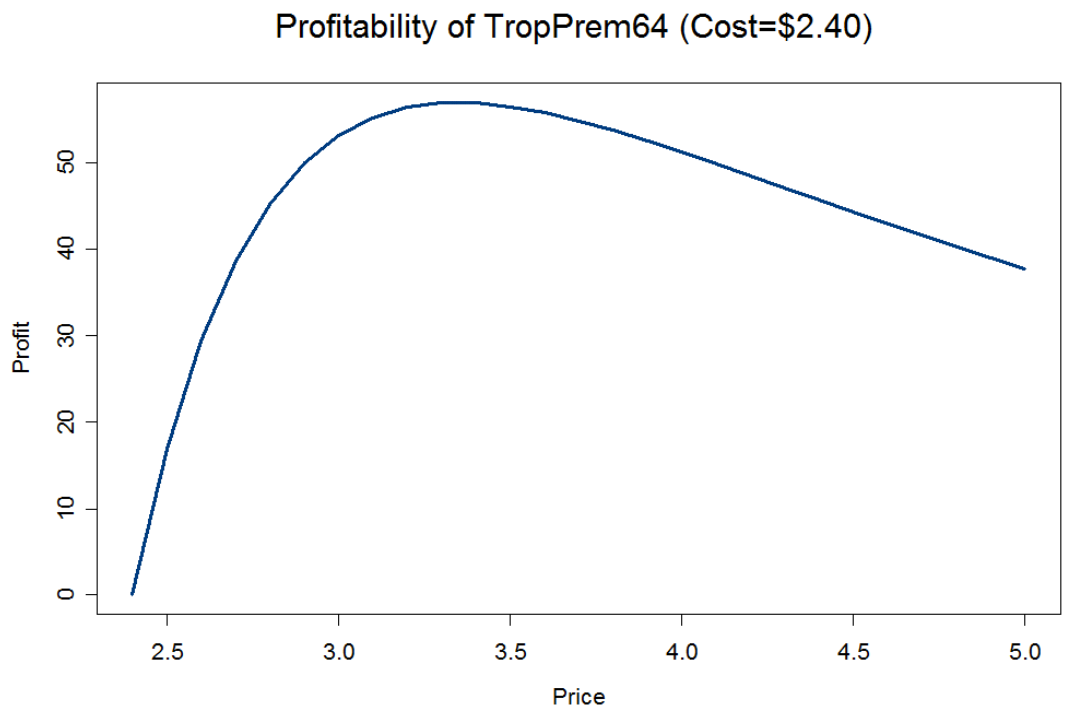
Creating informative priors from business rules

Empirical Example

Current State of Price Optimization


Movement vs Price of TropPrem64
Movement
Price
0
500
1000
1500
1.5
2.0
2.5
3.0

Optimal Product Pricing



Profits:

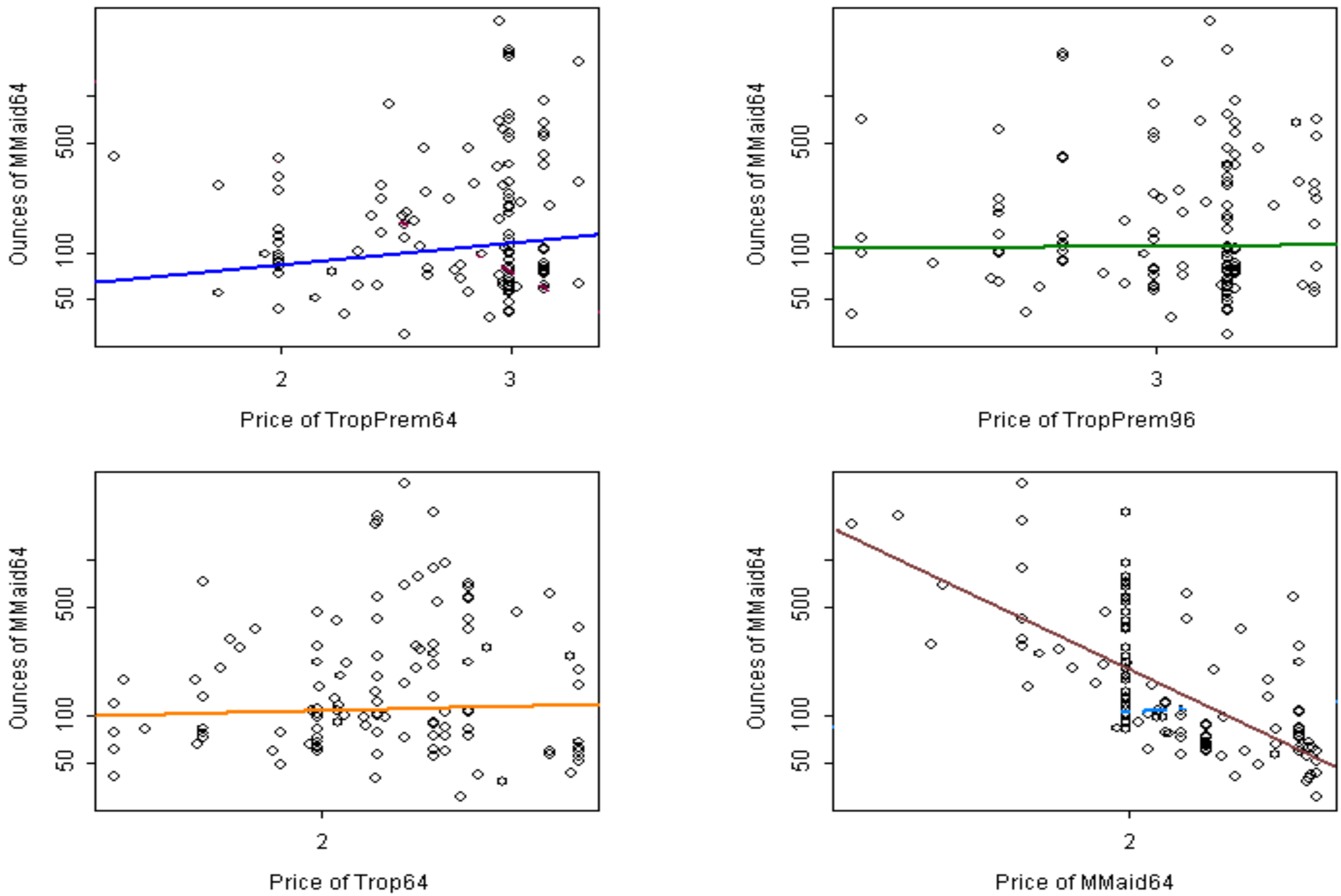
$$\Pi = (p - c)q$$

Optimal pricing rule:

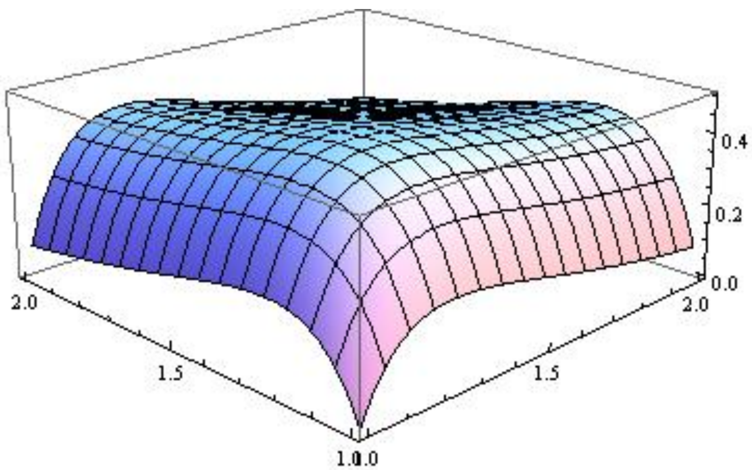
$$p^* = \frac{\beta}{\beta + 1} c$$

Where price elasticity measures demand responsiveness to price changes:

$$\beta = \frac{\partial q}{\partial p} \frac{p}{q} \approx \frac{\% \Delta q}{\% \Delta p}$$


Ounces of MMaid64
500
100
50
2
3
Price of TropPrem64
Ounces of MMaid64
500
100
50
3
Price of TropPrem96
Ounces of MMaid64
500
100
50
2
Price of Trop64
Ounces of MMaid64
500
100
50
2
Price of MMaid64

Optimal Product Line Pricing



Total Profits:

$$\Pi = \sum_{i=1}^{M} (p_i - c_i) q_i$$

Optimal pricing rule:

$$p_i^* = \frac{\beta_{ii}}{\beta_{ii} + 1 + \sum_{j \neq i} \mu_j \beta_{ji} \, s_j / s_i} c_i$$

Where cross price elasticities measure competitive effects:

$$\beta_{ij} = \frac{\partial q_i}{\partial p_i} \frac{p_i}{q_i} \approx \frac{\% \Delta q_i}{\% \Delta p_j}, \quad \mu_i = \frac{p_i - c_i}{p_i}, \; s_i = \frac{p_i q_i}{\sum_j p_j q_j}$$

Price Optimization in Practice

Huge growth of price optimization in practice for both retailers and manufacturers:

- IBM DemandTec, Dunnhumby's KSS Retail, Oracle, PROS, SAP Khimetrics, Vendavo, Vistaar Technologies, and Zilliant

Gartner Marketscope (2010) states that "price optimization technology will have a more direct impact on increasing revenue or margins than any other CRM technology".

- The Yankee Group estimated that more than one billion dollars would be spent on these systems in 2007

2012 Gartner study (Fletcher 2012) showed gains of 2 to 4% of revenue from using these systems.

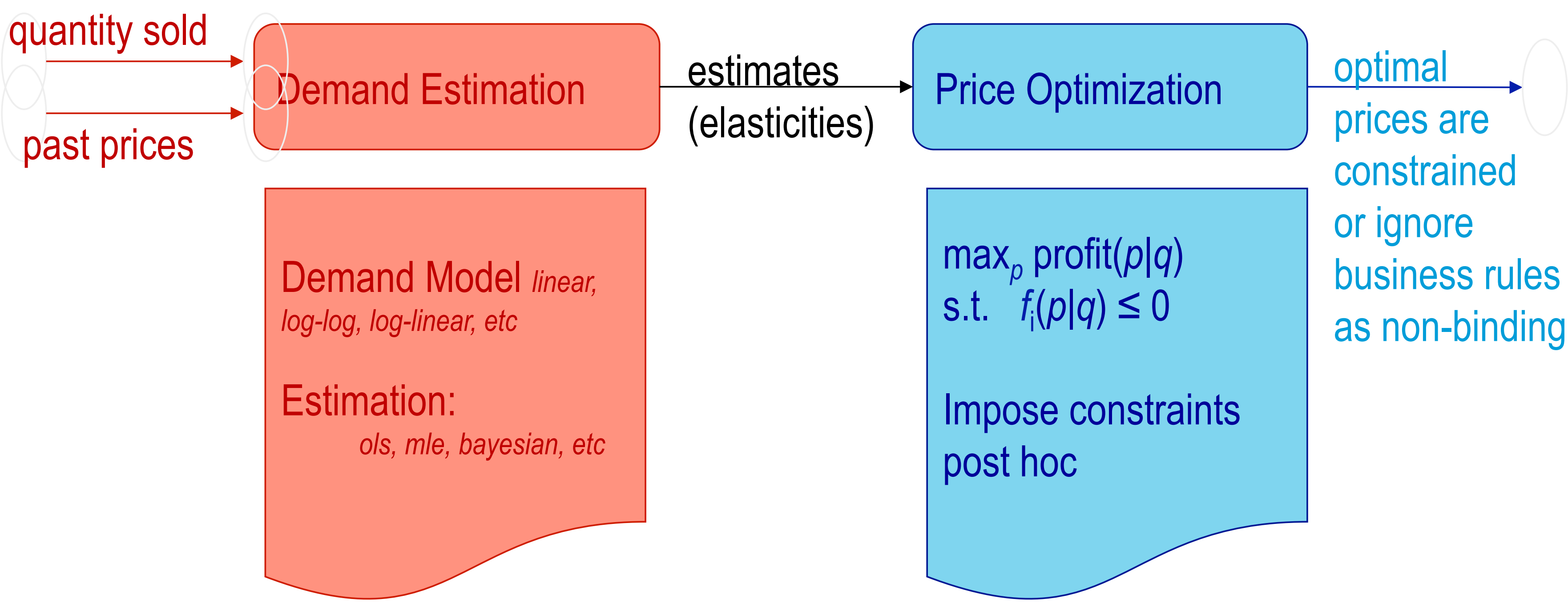
- Anecdotal reports suggest increases in gross margins in the range of 2-8%, retailers typically have gross margins of about 25% and have annual revenues of $2.5 trillion.
- Suggests benefit for retailers alone would be between $12.5b and $50b annually.

Germann et al (2014) find retailers benefit from deploying customer analytics more than firms in most inudstries

- Compred with Media Entertainment, Energy, Insurance, Telecom, Hospitality, Banking

Traditional Modeling Process

The multi-product pricing decision process:



Business Rules as Prior Information

Business Rules

Current pricing solutions frequently implement constraints that reflect “business rules”, which codify manager knowledge:

- Allowed number and frequency of markdowns (e.g., at least a week between two consecutive markdowns)
- Min-max discount levels or maximum lifetime discount
- Minimum number of weeks before an initial markdown can occur
- Types of markdowns allowed (e.g., 10%, 25%, …) or the permissible set of prices
- The “family” of items that must be marked down together

Source: Elmaghraby and Keskinocak (2003; Management Science)

Why use business rules?

Answer: to “improve” the pricing solution and find a better one than would be afforded without these constraints

Examples:

- Strategic decision (Elmaghraby and Keskinocak 2003)
- Ensure desired positioning of the product (Hawtin 2002)

DemandTec's Rule Relaxation Approach

1. Group price advance or decline rules. The user sets a maximum weighted group price advance or decline to 10%.

2. Size pricing rules. The user goes with the default that larger items cost less per equivalent unit than smaller identical items.

3. Brand pricing rules. For soft drinks, the user designates the price of brand A is never less than the price of brand B. For juices the user designates that brand C is always greater than Brand D.

4. Unit pricing rules. The user goes with the default that the overall price of larger items is greater than the overall price of smaller identical items.

5. Competition rules. The user designates that all prices must be at least 10% less than the prices of the same items sold by competitor X and are within 2% of the prices of the same items sold by competitor Y.

6. Line price rules. The user designates that different flavors of the same item are priced the same.

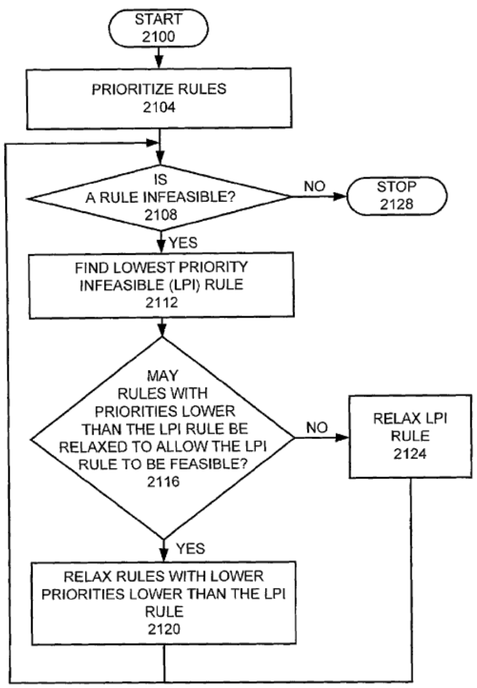


FIG. 21

Source: Neal et al. (2010), Patent #7617119

Examples of Business Rules in Academic Research

- Corstjens and Doyle (1981) use constraints to represent store capacity, availability and control
- Montgomery (1997) constrains the optimal price subject to an average price and/or total revenue constraint
- Subramanian and Sherali (2010) ensure that new solution does not deviate too far from current
- Deng and Yano (2006) bound price ranges to ensure reasonable solutions
- Bitran and Modschein (1997) and Feng and Gallego (1995) allow only a fixed number of price changes
- Gallego and van Ryzin (1994) prices have to be chosen from a discrete set of possible prices
- Reibstein and Gatignon (1984) prices of smaller sizes cannot be greater than the price of larger sizes
- Cohen et al (2015) uses rules to avoid promotion wearout, 9¢ price endings, and constraints on the number and timing of promotions.

Conjecture

Business rules and the constraints on the pricing solution they imply are a very popular feature of current pricing solutions

The problem is that this is a non-Bayesian solution since constraints represent a priori information about the solution (either to compensate for model weaknesses or to direct information about the solution)

A better approach (from a decision theoretic standpoint) is to follow a consistently Bayesian framework if we truly wish to have an optimal decision.

Business Rules constitute Prior Information

Any statements or constraints on optimal prices implicitly constitute prior beliefs

- The manager holds these beliefs even before looking at the data

Optimal prices are functions of price elasticities

Therefore, manager's prior beliefs about optimal prices imply priors on the parameter space

Prior beliefs should be represented as prior information if we follow Bayesian principles

Therefore, business rules should be employed a priori

Example

Consider the following situation:

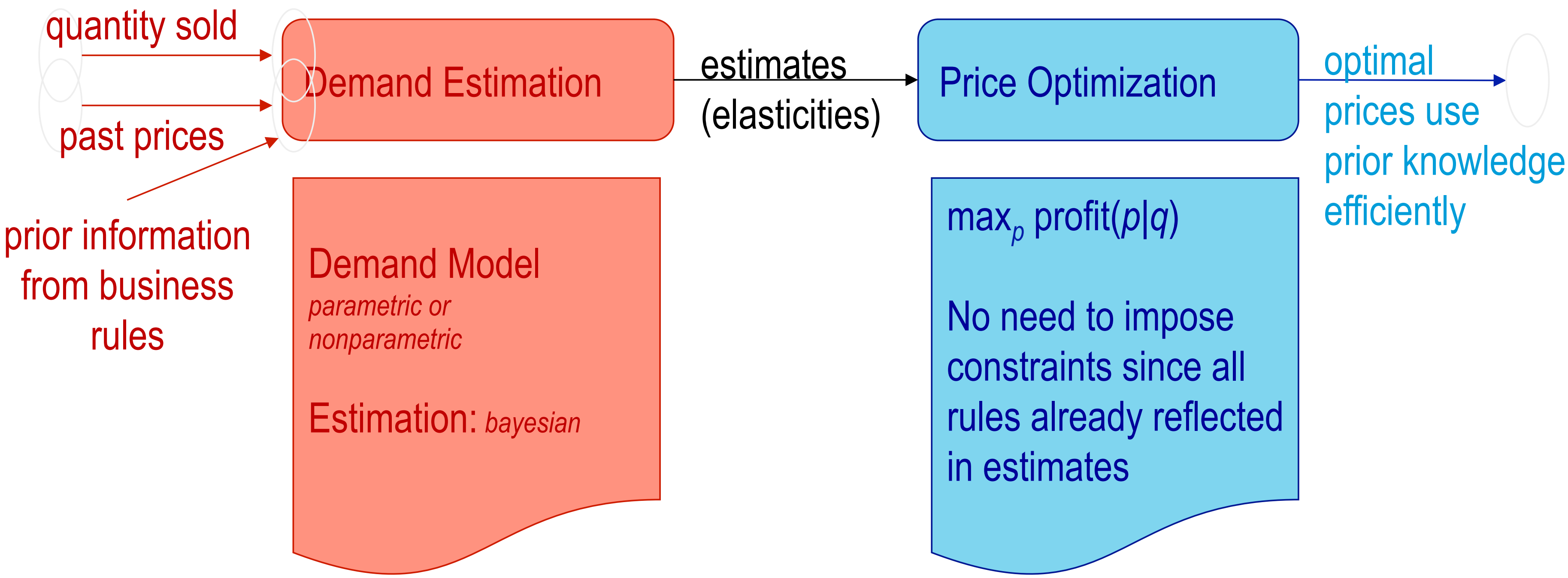
- A pricing manager receives the output from a pricing optimization solution, and then remarks that this solution doesn't look right—the optimal price should not be more than 5% higher.

There are two possibilities:

- The pricing manager has independent information from the data/ model (e.g., was not involved in the analysis) and is trying to combine two sources of information
- The pricing manager has dependent information from the data/ model and is trying to influence the posterior solution to conform to his prior beliefs. However, the natural sequence is inverted – posterior/data → prior

Proposed Modeling Process

The multi-product pricing decision process:



Creating Informative Priors from Business Rules

Practical Problem

Bayesian models force analysts to formulate priors in terms of conjugate (or non-conjugate) priors *on the parameters* which represent all known information.

The manager in the previous case possesses an assessment about optimal price—a marginal property of the model—not the parameters of the model.

Our proposal is to infer a prior from the marginal distribution by inverting the manager's assessment about the optimal price into an implied prior.

Log-Linear Demand Example

Suppose quantity demand follows a log-linear model, where q is quantity and p is price:

$$\log(q) = \alpha + \beta \log(p) + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2)$$

The optimal price (conditioned upon parameter) and the cost (c):

$$p^* = f(\beta) = \frac{\beta}{1+\beta} c$$

And we have a business rule:

$$c \le p^* \le \tau$$

Change of Variable Solution

More generally we can think of our optimal price as a function of the price elasticity:

$$p^* = \frac{\beta}{1+\beta} \cdot c = f(\beta) \cdot c, \text{ where } f(\beta) = \frac{\beta}{1+\beta}$$

We can transform between the optimal price space and the price elasticity space using inverse functions:

$$\beta = f^{-1}\left(\frac{p^*}{c}\right) = \frac{p^*}{c-p^*}, \text{ where } f^{-1}(x) = \frac{x}{1-x}$$

The implied prior can be computed from a change of variables approach, which gives the distribution of optimal prices:

$$p_p(p^*) = p_\beta\left(f^{-1}\left(\frac{p^*}{c}\right)\right) \cdot \left|\frac{df^{-1}}{dp^*}\right| = p_\beta\left(\frac{p^*}{c-p^*}\right) \cdot \frac{c^2}{(c-p^*)^2}$$

Decision Rules

Bayesian Rule maximizes the posterior expectation of profits.

$$\tilde{p} = \arg\max_{p} E\left[\pi\left(p,\beta\right) | \mathscr{D}, \mathscr{R} : p^{*} \in (c,\tau)\right]$$

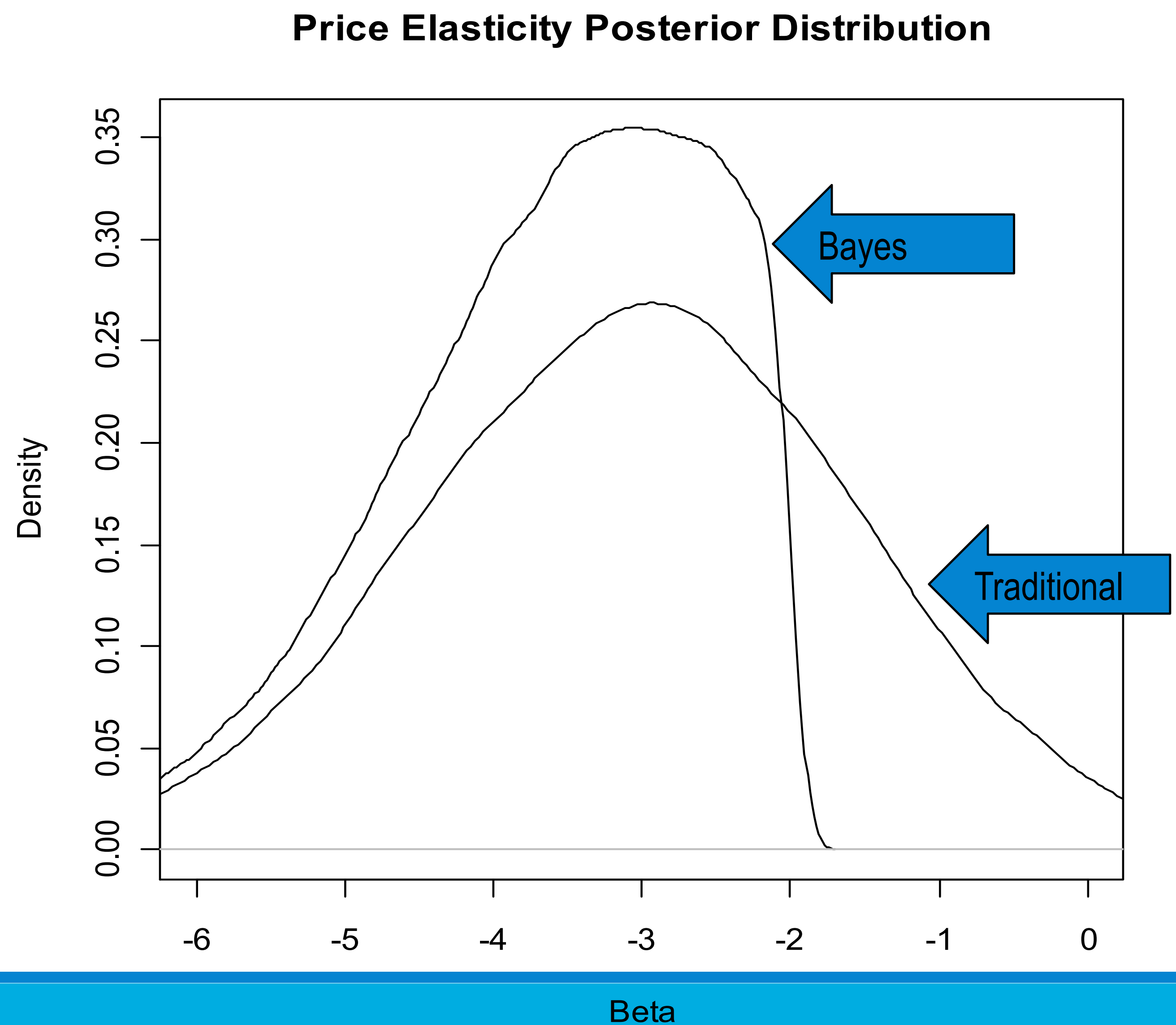
The traditional approach:

$$\tilde{p} = \min\left(\arg\max_{p} E\left[\pi\left(p,\beta\right) | \mathscr{D}\right], \tau\right)$$

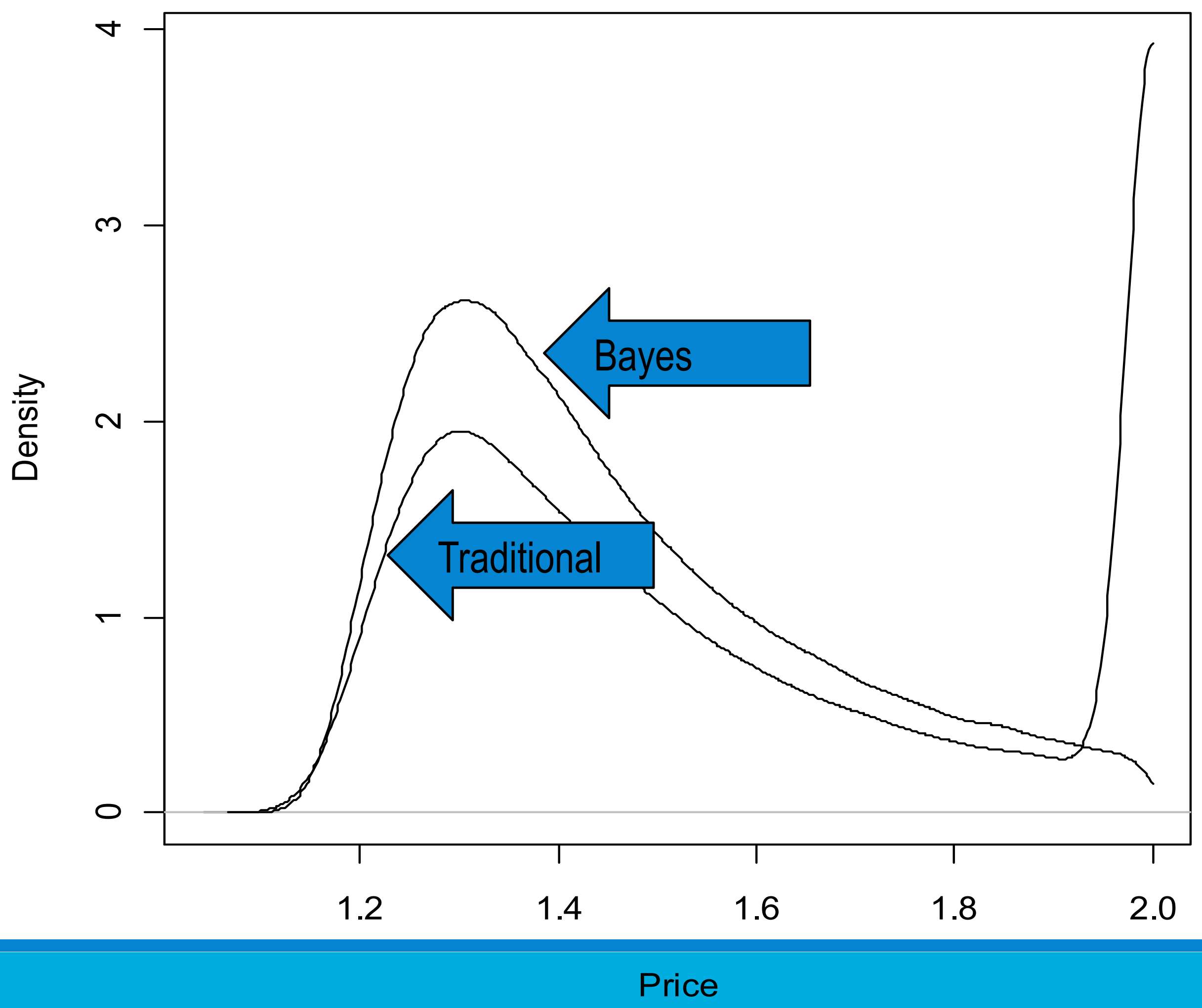
The difference between the two is the Bayesian rule is unconstrained – it already incorporates the information, while the traditional approach introduces the rule post hoc.

Caution: Prices used in three distinct contexts:

- Observations (Data), Optimal Price (Distribution), Price from Decision Rule (Point)


Price Elasticity Posterior Distribution
Density
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
-6
-5
-4
-3
-2
-1
0
Bayes
Traditional
Beta

Optimal Price Posterior Distribution



Example Results:
Elasticity and Optimal Price Estimates

Rule	Price Elasticity		Optimal Price	
Traditional	-3.00	(1.00)	1.67	
Bayesian	-3.29	(0.79)	1.49	(0.18)

Weak Prior on Elasticity, optimal price < $2.00

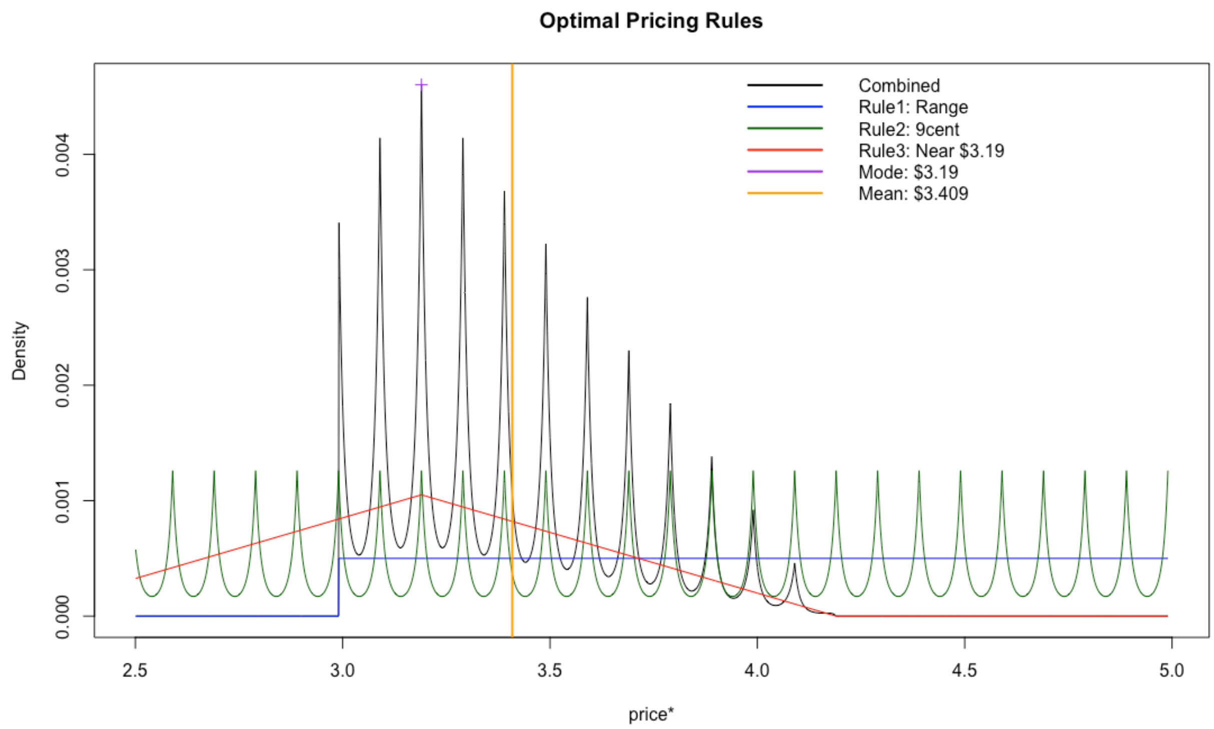
Example Results:
Elasticity and Optimal Price Estimates

Rule	Price Elasticity		Optimal Price	
Traditional	-3.00	(0.71)	1.50	
Bayesian	-3.57	(0.43)	1.40	(0.06)

Strong Prior on Elasticity, optimal price < $1.50

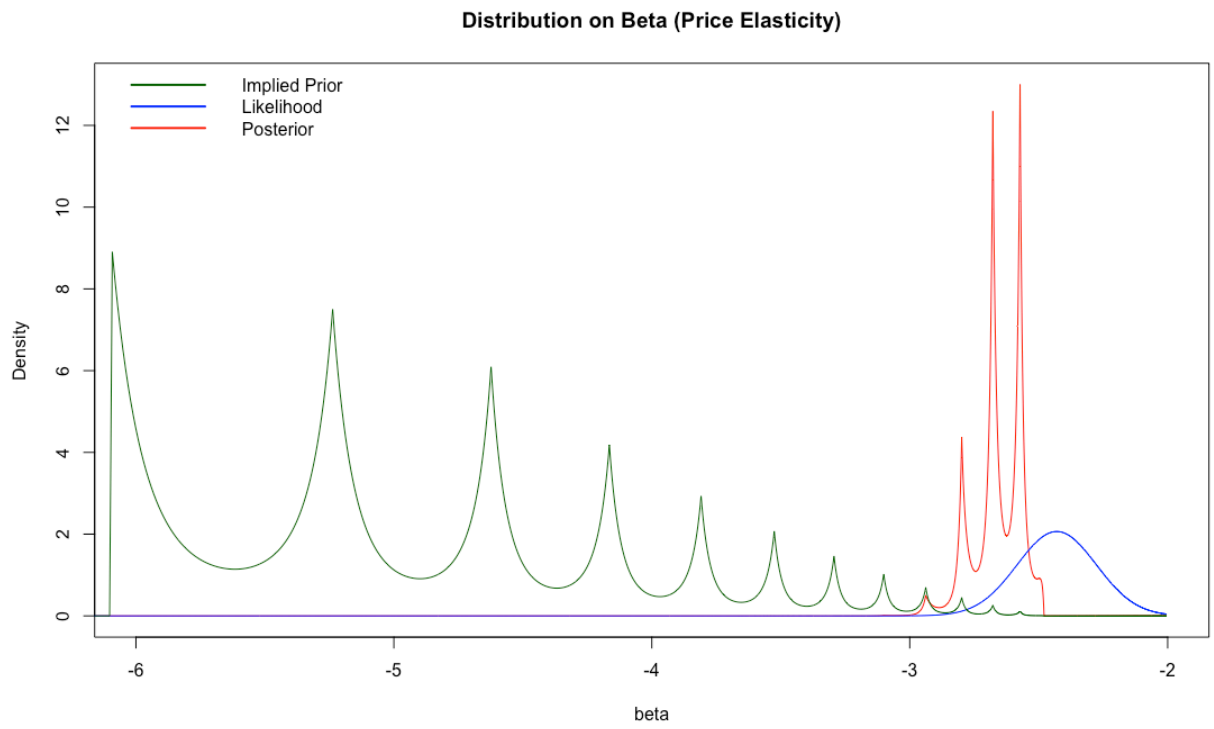
Example 2: Optimal Price Prior with 3 Rules

1) Range ($2.99,$4.99), 2) 9¢ endings, 3) Near $3.19



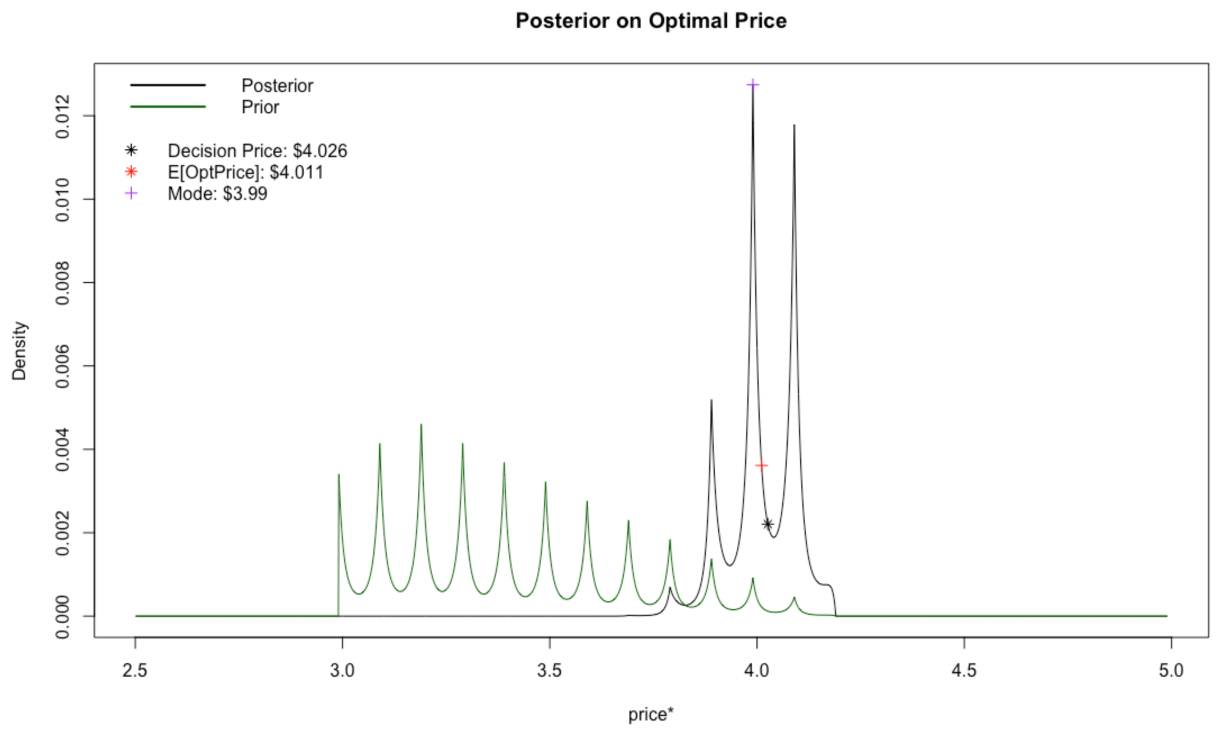
Example 2: Optimal Price Prior with 3 Rules

1) Range ($2.99,$4.99), 2) 9¢ endings, 3) Near $3.19



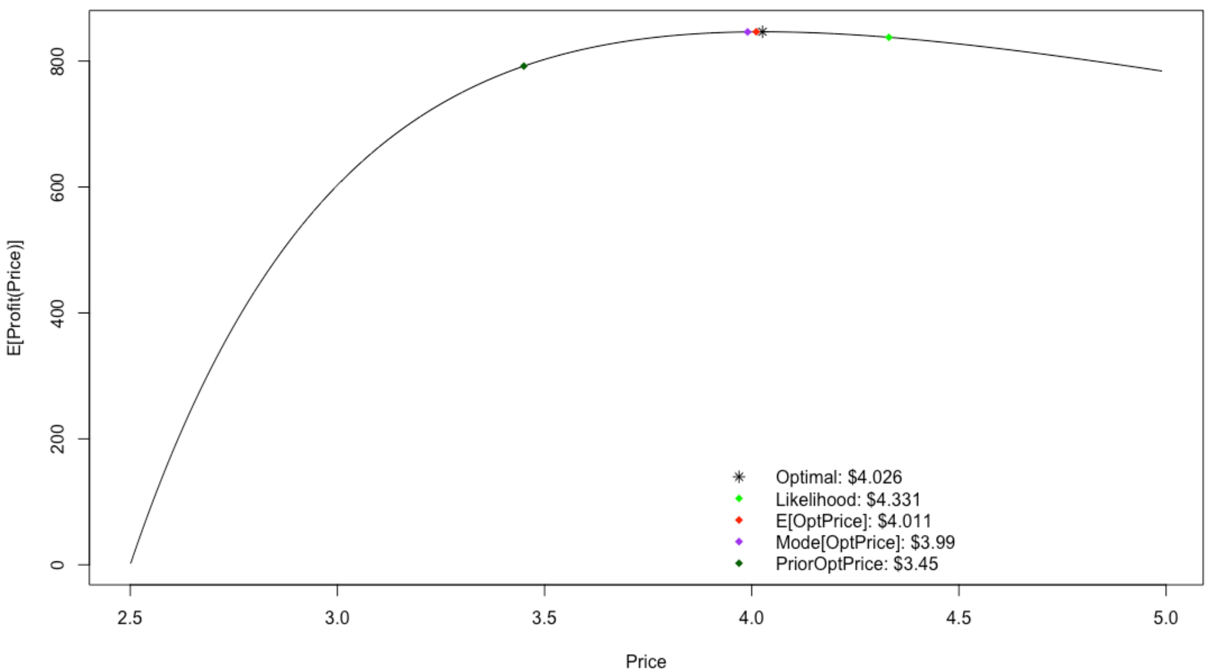
Example 2: Optimal Price Prior with 3 Rules

1) 9¢ endings, 2) Near \$2.99, 3) Range (\$2.99,\$4.99)



Example 2: Optimal Price Prior with 3 Rules

1) 9¢ endings, 2) Near \$2.99, 3) Range (\$2.99,\$4.99)



Summary

The traditional approach splits the problem into an estimation and inference scheme.

- Business rules are used in an ex-post manner, and if they are not binding they are ignored.

The inferential approach is Bayesian during estimation.

- Here the optimal price is constrained but the parameter estimates are not changed.

The Bayesian Decision Theoretic approach combines estimation and inference into a single task.

- The estimates from both are consistent and efficient. Information from business rules are used a priori and optimal prices are unconstrained.

Substantial and practical differences amongst the approaches.

Empirical Example

Multivariate Pricing Problem

System of demand equations over M products:

$$\ln(q_{it}) = \alpha_i + \sum_{j=1}^{N} \eta_{ij} p_{jt} + \xi_i f_{it} + \psi_i d_{it} + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, \sigma_i^2)$$

The manager wishes to optimize total profits when faced with known variables costs:

$$\Pi = \sum_i (p_i - c_i) \mathrm{E}\left[q_i\right]$$

Optimal price margins solve the first order condition:

$$\mathbf{m} = -diag(\mathbf{w})\left(\mathrm{B}'\right)^{-1} \mathbf{w} = -diag(\mathbf{r})\left(\mathrm{B}'\right)^{-1} \mathbf{r}$$

Business Rules

Individual price bounds

- Optimal prices are within 20% of current prices
- Alternative: prices above cost and lower than cost x 150%

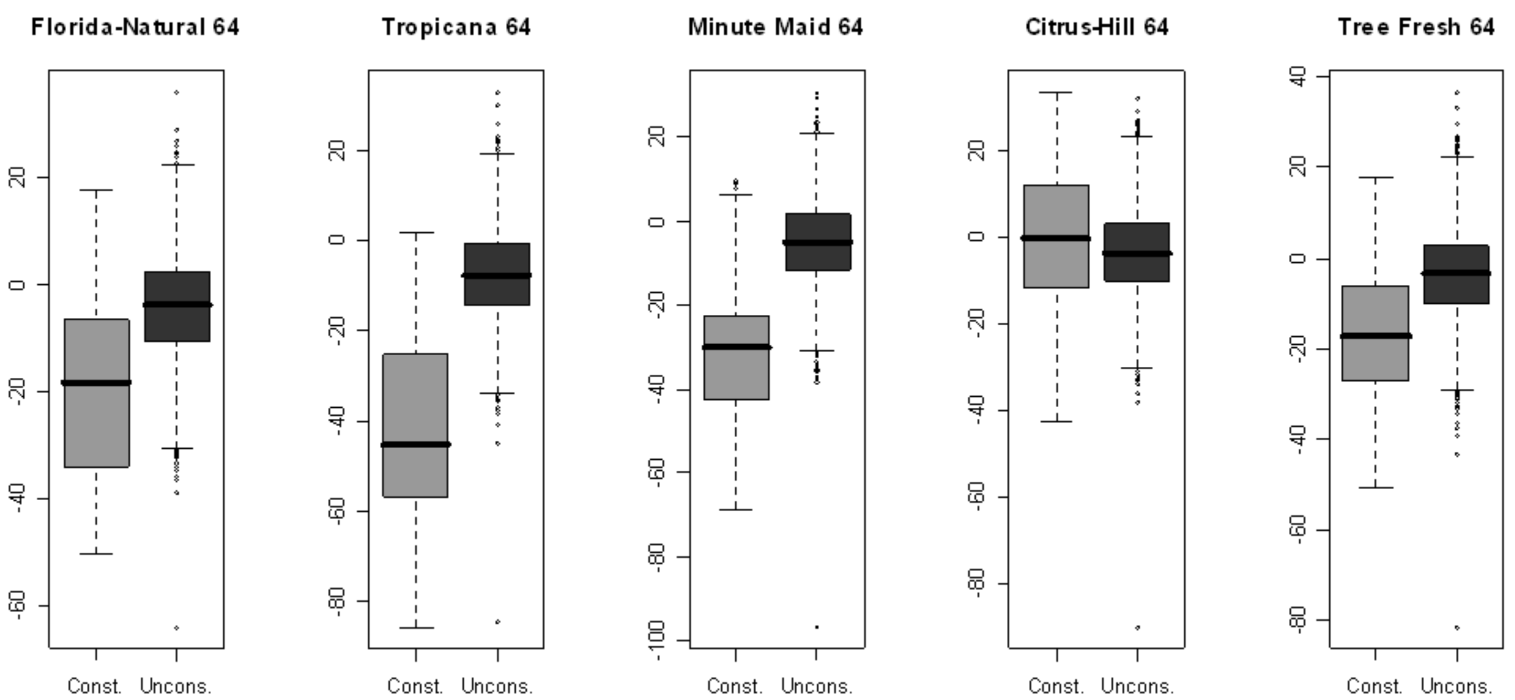
$$\left\{ p_{0i}^{*} \leq p_{i}^{*} \leq p_{1i}^{*} \right\}$$

Enforce price-quality tiers

- High-quality tier are greater than national brand prices
- National brand prices are greater than store brand prices

$$\max\left(\mathbf{p}_{A}^{*}\right) \leq \min\left(\mathbf{p}_{B}^{*}\right), \max\left(\mathbf{p}_{B}^{*}\right) \leq \min\left(\mathbf{p}_{C}^{*}\right)$$

Posterior Price Coefficients



Optimal Prices

Tier	Product	Traditional	Bayesian	Change
Premium	TropPrem64	.0526 (.0033)	.0515 (.0046)	-2.1%
National	TropReg64	.0408 (.0027)	.0379 (.0033)	-7.1%
National	MinMaid64	.0409 (.0028)	.0360 (.0029)	-12.0%
Store	Dom64	.0321 (.0022)	.0276 (.0029)	-14.0%
	Total Profits:	\$11,074 (2,372)	\$7,073 (1,135)	-36.1%

Findings

Treating business rules as prior information has a huge impact on

- Parameter estimates
- Optimal prices
- Expected profitability

Conclusions

Conclusion

Current optimal pricing practitioners and researchers are introducing information in an ad hoc manner by relying upon “business rules” or constraints.

An appropriate Bayesian method is correct and efficient and should avoid ad hoc “corrections” to the posterior.

This approach is easy and could lead to better decision support systems that reflect “expert” knowledge efficiently.

Further Results

We have embedded the optimization problem within a MCMC algorithm, which gives a general approach for optimization under uncertainty

- We relax the certainty of the business rules (perhaps the data can inform whether the business rule is correct)
- Sometimes business rules are used to test the model (e.g., if the model gives strange solutions then it is wrong)
- Relax the assumption about functional form and can learn it from the data using nonparametric techniques