Masters Dissertation

Understanding and Supporting Pricing Decisions using Multicriteria Decision Analysis:
An Application to Antique Silver in South Africa

JED STEPHENS

Supervisor: Emeritus Professor T.J. Stewart

Statistical Sciences
University of Cape Town

Minor Dissertation presented in partial fulfilment for the degree of:
Master of Science (specialising in Advanced Analytics)

Science Faculty Ethics Approval: FSREC 32 - 2019

February, 2020

Abstract

This dissertation presents an application of multicriteria decision analysis to understand and support pricing decisions in fields where goods are unique and described by their characteristics. The specific application area of this research is antique silver objects, where a complete iteration of the multicritia decision process is performed. This includes two problem structurings using SODA which provide rich detail into this application area. Multi-attribute additive models are constructed, with attribute partial value functions elicited using different methods: directly (bisection methods), indirectly (MACBETH and linear interpolation) and with discrete choice experiments. The applicability and advantages of each method is discussed. Additionally, an open source R package to implement the design of discrete choice experiments is created. The multi-attribute models provide key insights into decision maker's reasoning for price; and contrasting different decision maker's models explains the market. A risk adverse relationship between multicriteria model score and price is characterised and various inverse utility functions investigated. Two decision support systems are fully developed to address the needs of Cape silver decision makers in South Africa.

Keywords— MCDA, Pricing, Antique Silver, Cape silver, SODA, MACBETH, Utility Functions, Decision Support Systems

Dedicated to the antique silver dealers who gave many hours of their time. Nothing would exist without your participation.

Acknowledgements

To Professor Stewart for his supervision, good humour, critique and encouragement. Your comments and suggestions were insightful and improved this work.

To Alexes for her love, support and way of making me smile each day. Thanks for comments and proof-reading of earlier drafts.

To my family, Blake, Carla, Peter, Penney and Roux for always being interested, loving and caring. Special thanks to Carla for comments on earlier drafts.

To the Potter Foundation, in appreciation of my award of the Potter Fellowship, and for supporting this work as well as the following three during my MSc:

Nattrass, N., Stephens, J. and Loubser, J.J., 2019. Animal welfare and ecology in the contested ethics of rodent control in Cape Town. *Journal of Urban Ecology*, 5(1), p.juz008.

Nattrass, N., Conradie, B., Stephens, J. and Drouilly, M., 2019. Culling recolonizing mesopredators increases livestock losses: Evidence from the South African Karoo. *Ambio*, pp.1-10.

Conradie, B., Piesse, J. and Stephens, J., 2019. The changing environment: Efficiency, vulnerability and changes in land use in the South African Karoo, 2012–2014. *Environmental Development*, p.100453.

Abbreviations

DCE: **D**iscrete **C**hoice **E**xperiment

DSS: **D**ecision **S**upport **S**ystem

DM: **D**ecision **M**aker

MCDA: **M**ulti**c**riteria **D**ecision **A**nalysis

MCDM: **M**ulti**c**riteria **D**ecision **M**odel

PSM: **P**roblem **S**tructuring **M**ethod

Contents

Chapter 1

Introduction

Broadly this work investigates what we can learn about the use of multicriteria decision aids to assist in the pricing of goods which are non-reproducible (i.e. unique) and desirable, as well as describable by their characteristics. The aim is to understand more about the market of antique silver objects: in particular how the describable features of an antique silver object come together to determine a perception score (or utility) – that score ultimately being expressed as (or converted into) a monetary price. Currently attempts to establish a price for these objects resemble something of an unstructured approach to multicriteria decision analysis (MCDA). Marttunen et al. (2017, 1) describes MCDA as "a collection of systematic approaches developed specifically to support the systematic evaluation ... in terms of multiple and often conflicting objectives." This study is the first (known to this author) application of MCDA not only to antiques, but to the decorative arts. It examines the usefulness of such an approach.

In the late 1880's a distinction was made between the fine arts and the decorative arts (Arts & Crafts Exhibition Society, 1888; The Editors of Encyclopaedia Britannica, 2019). Silversmithing came to be considered a decorative art – as almost all pieces of wrought silver have some usefulness[1] – yet for historical reasons it is important to understand that until this distinction "a silversmith could be considered the equal of any painter or sculptor – and indeed may have been trained in any of these disciplines" (The Silver Society, 2019). Despite this distinction, silver objects straddle an awkward divide between the two arts: especially because the aesthetics of a silver object is an important theme.

The situation is further complicated because this is a study of antique silver objects and such objects have 'worth' not only in the context of the characteristics of the silver object, but in the characteristics which make it antique – a working definition of an antique being: "a primarily handcrafted object of rarity and beauty that, by means of its associated provenance and its agedness as recognised by means of its style and material endurance, has the capacity to generate and preserve for us the image of a world now past" (Rosenstein, 2009, 14).

In the decorative arts the valuation of jewellery has become increasingly codified in something resembling a MCDA system. As an example, the Independent Coloured Stone Laboratory (ICSL) provides a grading system which allows trained professionals to form the same opinion on the quality of cut gemstones – this score then dictates the cost per carat of stone. Similar specialised reports are compiled for diamonds, such as the Gemologist Institute of America (GIA) report. The valuation of jewellery (such as for retail or insurance purposes) is then worked up from its inputs: the metal and gemstones as well as an estimate of labour time and other factors. While this could not be called a MCDA price valuation system, there are at least some flavours of it – which is more than can be said for the price valuation of antique silver objects.

[1] In contrast a fine art is considered an expression that uses skill or imagination in the creation of a purely aesthetic object (The Editors of Encyclopaedia Britannica, 2018).

In the fine arts there is no known example of the application of MCDA models to the evaluation, let alone pricing, of objects. Nonetheless, there is certainly the need. The situation such as Pogrebin (2019) describes, where museum collections require consolidation, would benefit from a MCDA process which makes the decisions undertaken by curators both explicit and understandable by the general public. It is a portfolio problematique (Belton and Stewart, 2010, 15). Curators could also compare their different MCDA models in order to understand what different institutions value. The MCDA could further serve as a basis of a recurring decision or evaluation (Belton and Stewart, 2010, 30) for the purchase of new artworks under financial constraints.

To return to this work, the research will be conducted using in-depth interviews with experts (either retail shop owners or auctioneers) whose businesses focus on antique silver. That is to mean antique silver typically makes up a majority proportion of their business. The interviews take part in the greater Cape Town area, South Africa. Each interview was typically multi-visit, which consisted of a first session with formal problem structuring and then a later in-depth MCDA process[2] and a questionnaire[3]. Henceforth these experts (who were respondents) are referred to as the decision makers (DMs). Questionnaire results characterise the DMs as having both a large reliance on their intuition and yet still considering their choices systematic[4]. A position which Kahneman (2011) would probably consider paradoxical: with these current decisions being examples of "thinking fast" activities[5].

How then indeed have objects of antique silver been priced? Historically objects of antique silver were priced with a currency per weight measure (Waldron, 2001). This was the case up until the late 1970s when the rise of 'price guides' such as Miller and Miller (1979) occurred. These price guides feature auction style-like estimates for objects (including, but not limited to antique silver) from premier auction houses. The price estimate is given as a region bounded by a low and high price, where the region is intended to reflect some uncertainty about the price[6]. This decision support tool will be referred to as an 'estimated price guide'.

Another source of information is past auction results. For other fields, such as the fine arts, there are decision support tools (as an example Art Price (2019)) which aggregate these historical records into a searchable database. There is no such equivalent for objects of antique silver. Auction historical records do not always reflect retail prices[7] and, notwithstanding this, the notion of a "market" price is difficult when the goods are defined by their characteristics and are unique. These historical records form a type of decision support tool that will be referred to as a 'historical price guide'.

Since the historic transition (away from a price per gram) to giving a price for the overall object, the role of the scrap (or metal) price of silver comprising the object has been somewhat obscured. Draper et al. (2018); Stephens (2017) both found (not unsurprisingly) that the scrap price of silver

[2]The application of multiattribute value theory by constructing an additive model with attribute partial value functions elicited using different methods: directly (bisection methods), indirectly (MACBETH and linear interpolation) and with discrete choice experiments

[3]The questionnaires covered background information as well as world views of the antique trade.

[4]The respondents were asked two Likert scale questions. The first was "when pricing an object of antique silver to what extend do you rely on your intuition (about what the price is)" with the response options being: highly reliant, reliant, unreliant and highly unreliant. The second question was "how would you describe your approach when pricing an object of antique silver" with the response options: highly systematic, systematic, unsystematic and highly unsystematic.

[5]More than one DM explicitly explained that they needed to make pricing decisions quickly especially when buying in pieces as trading stock. As the supply of antique silver is dependent on customers selling their goods to a DM, the DM has to provide a purchase price while they are inspecting the object (often for the first time) in front of the client. In determining a purchase price, the DM would nonetheless need to have some vague idea of a retail price. This work is focused on retail prices.

[6]The source of this uncertainty is intentionally unclear: it offers some opportunity to justify to a bidder a payment higher than the mean estimate point as well as reflects auction room dynamics.

[7]Some auctioneers position themselves as retail auctioneers while others as wholesalers. There is no data available to separate auction houses, even by this coarse level of distinction.

creates a binding lower bound (floor) for the price of antique silver on auction in England and South Africa respectively. This is unlike in the fine arts where the inputs are relatively negligible (e.g. oil paints) to the subsequent reselling of a painting. Interestingly the premium mark-up above the scrap price is still largely correlated with the fine silver mass[8], see Figure 3.2 (page 11). The role of scrap price re-features in trying to decide on a preferential independence for the scale of any MCDA.

With the current decision support tools the methodology is simple: the DM matches the intended object as closely as possible with available records. If multiple records match the DM will attempt a weighted interpolate between them. This is not trivial – not only because the one source may be historic and the other an estimate – but because matches are usually partial at best. Some records will match on silversmith, some on the style and period, some on the size of the object – and then the DM must decide how to weight the interpolation. Notwithstanding these difficulties the current decision support tools are loved by the DMs and will play an interesting role in the interactions with the DMs.

This dissertation seeks to introduce a new method of price estimation. The appropriateness of MCDA is clear: the nature of pricing is a recurring decision or evaluation over an ever-changing set of actions (alternatives) most closely akin to the design problematique (Belton and Stewart, 2010, 15, 30). A pricing decision is firmly planted in what Keeney (1996) described as "value-focused thinking". In support of this notion is the following result obtained from the questionnaire which allowed the DMs to move a continuous sliding 100 point scale rating their pricing motives between personal (-50), neutral (0) and objective (+50). Tuckey's Five-Number summary thereof is -37, -33, -30, 0, 3 which reflects how DMs perceive estimating price based on personal value[9]. This personal value determines some score (a large focus of problem structuring was what score) and that score has some relation to price. A risk adverse relationship is characterised in Chapter 7.

The study seeks to achieve two major outcomes: firstly to better understand DMs thought process to price his/her antique silver object, reflecting the personal nature of the problem. Secondly, to use these insights to provide help with pricing decisions by creating decision support systems. A possible exploration point for future experiments is the use of discrete choice experiments (DCEs) as an indirect value elicitation method. The original motivation for DCEs stems from wanting to investigate the effects of mass and its role in price. The methodology is divided into problem structuring and its results (Section 4.1), MCDA methods (Section 4.2) and methods around the design and analysis of DCEs (Section 5.3). In order to implement DCE designs, the R package `ExpertChoice`[10] was written as a research output of this work and is available on CRAN. Appendix C provides more details about the package, as well as the package vignettes.

The results have two temporal parts, early multicriteria decision modelling (Chapter 6) and a later modelling of a focused single origin market (Chapter 7) for Cape silver. The need for these two separate chapters reflect how this analysis evolved during implementation. Together they show a complete iteration of the MCDA process, including a reformulation of the original problem structuring in the latter section. The earlier Chapter 6 also captures an important theme which runs throughout this work: the interactions and expectations of the DMs. The interactions in this study – unlike in most MCDA applications where the DM requests the assistance of the analyst – were unsolicited. The role this played in the pre-interview process and in the expectation of reciprocation for their time – Section 6.6 provides more detail. The result of meeting these expectations is the web based decision support system `CapeSilversmiths`, http://www.capesilversmiths.co.za, which provides historical records from South African auctioneers for Cape silver.

[8]Fine silver mass is the mass of pure silver in an object.

[9]Two corroborating questions support this finding. The DMs were also asked if they would overprice objects that they personally like and separately if they would underprice objects that they personally dislike. All (n = 7) but one of the DMs stated that they would either underprice or overprice objects that they personally like or dislike.

[10]The author acknowledges that there is commercial software with a similar name and that no relation was intended.

Pricing decisions are onerous, yet important. A multicriteria model can assist a DM to arrive at replicable (reproducible) pricing decisions. This model is typically flexible, understandable and rooted in the value assessments from the DM. As such, it avoids many of the problems associated with existing decision aids. An important part of any multicriteria modelling process is ensuring that the results can be used by the DM. To achieve this, a knowledge-based decision support system is created demonstrating how one of the DM's multicriteria model could be used. This decision support system, `CapeSilverDecisions`, is available online at https://capesilverdecisions.herokuapp.com/. The construction and analysis of decision support systems is the focus of Chapter 8.

Finally, the conclusion (Chapter 9) reflects on the role of this study in the arts, provides further research opportunities and highlights its contributions to the literature.

Chapter 2

Antique Silver Objects

The purpose of this short chapter is to give a working introduction to antique silver objects. The term origin is used to denote that these objects originated from a particular locality, examples would be: British, Cape and Irish. Chapter 7 focusses on multicriteria decision modelling with Cape (origin) silver. Since some foundational details (about this origin) are more relevant to this chapter, they are described here.

Many details of silver objects have already been given: that silversmithing is a decorative art and, as such, form follows function for silver objects. As Pickford (2010) explains silver has been wrought into objects for centuries. Examples of silver objects surviving to the present day are more common from about 1500CE onwards. Around this time a practise in Britain which would spread to the rest of Europe became more common: the application of a punched (using a hammer) symbol or mark to the silver. (Figure 2.1 gives an example.) These marks have evolved various uses and ubiquitous presence: the complexity of marking process becoming synonymous with silversmithing.

The earliest (and still most common) is the maker's mark[1]. A silversmith would typically have assistants who would craft silver only for it to be marked by the master silversmith. To address this, some literature suggest the more inclusive term of "sponsor's mark" rather than maker's mark. Maker's mark is the term used in this work because it is still understood ubiquitously. The defining feature of an object which has a maker's mark remains its master silversmith who would have been responsible for design, quality control and retail (either to the trade or the public) of the produced silver.

Marks applied by third parties (other than the maker) also evolved: some uses were to control standards (assay marks), tax (duty marks[2]) and track movement across country borders (import marks[3]) are the most common. The purity (the percentage of pure silver comprising the object) is usually legislated (on a country or regional basis) and enforced by a controlling authority called an assay office who validates that the object meets the minimum standard. Should the standard not be met the assay office confiscates the object (and destroys it). Every object a silversmith produces is required to be assayed. The assay mark is the special mark applied by an assay office recognising that it has been checked. Each country typically has different standards of purity legislated, but amongst collectors these are well known.

[1] In the English language it is usually in the form of the maker's initials. In other origins (especially France, Spain and China) it could be a pictorial symbol.

[2] A special mark applied during certain times to show that an additional tax was paid.

[3] A mark applied by an import office indicating that the silver object was imported (and that taxes were paid on the import).

In general all punched symbol marks (including maker's marks) are referred to as hallmarks. Figure 2.1 shows a set of hallmarks[4]. Excellent reference texts exist such as Miller (2017), Whetstone et al. (2010) and Niklewicz et al. (2016) that allow hallmarks (of any origin) to be interpreted. It is the marking of silver in this way which allows buyer and seller to be certain about many factors such as age, silver purity, maker and origin.

Figure 2.1: A set of Cape (origin) silver hallmarks. The mark on the far right is that of Johannes Marthinus Lotter (Active 1844 to 1879) i.e. his maker's mark. The other two marks appear like guild-hallmarks, but in fact are pseudo-marks. These imitation marks (which come in many variations) were applied by makers of Cape (origin) silver to suit the taste of European customers. Unlike British silver, objects of Cape (origin) silver were never assayed.

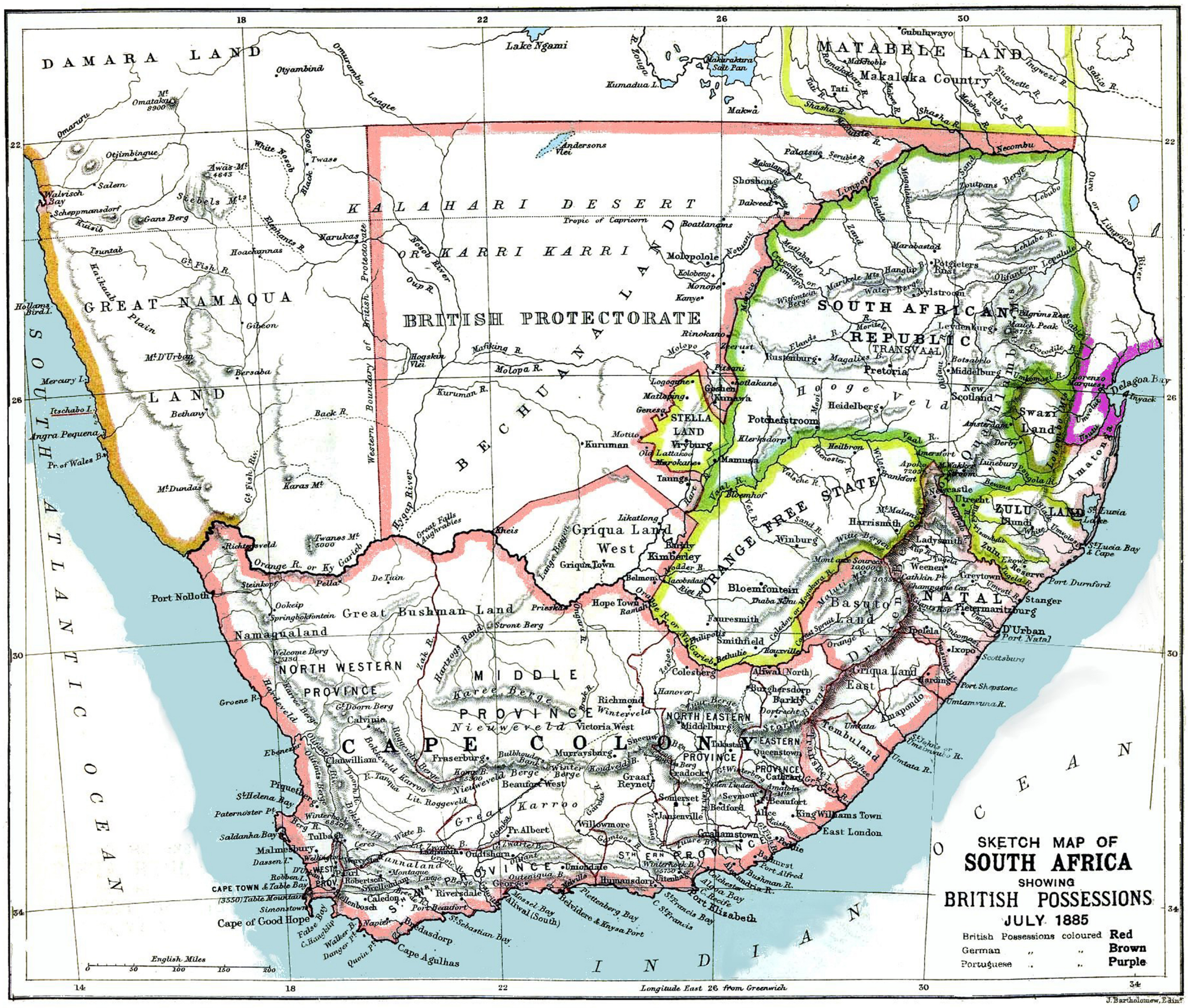
2.1 Cape Silver

South African wrought silver spans from the late 17th century until the present day. The map (Figure 2.2) shows South Africa in 1885 which is approximately the end time that silver made in South Africa would be considered Cape silver: named after the location in which it was made[5]. For Cape silver the authoritative study by Welz (1976) has a cut off date of the mid-19th century, the studies by Heller (1949, 1953) cut off in the year 1870 and Morrison (1936) cuts off in 1850. These end points capture a period of time when the economy of locally produced silver objects could not viably compete with British hallmarked silver. Furthermore because British hallmarked silver was already purchased in preference, narrowing cost margins brought an end to both demand and supply (Welz, 1976) of Cape silver. The last Cape silversmith was Jacobus Johannes Vos

[4]This term is distinctly English in origin and refers to the guild halls who were responsible for the assay process hence "hallmarking".

[5]Silver objects produced up until this period take the name of Cape silver after the region named the Cape (by the VOC) and later known as the Cape Colony (see Figure 2.2).

who in 1882 ceased to be an active silversmith finally ending the supply of locally produced silver. Acknowledgement of the comparative scarcity of Cape silver as well as the quaintness of style and form were at the heart of renewed collector interest (Heller, 1949). There are fifty-five distinctly identifiable silversmiths[6] presently. Cape silver is considered some, if not the rarest, of colonial silver. Output was limited by the availability of silver at the Cape[7] as well by the relatively few silversmiths who were active.



Scottish Geographical Magazine, 1885.

Figure 2.2: Map of South Africa in 1885. Source: Bartholomew, J., 1885, *Scottish Geographical Magazine.*

[6]Welz records additional marks of silversmiths some of which remain unidentified.

[7]Welz notes that old silver was often the source of silver. The mining of silver was limited.

Chapter 3

Mass and its role in price

With objects of silver there is a longstanding historical link between its material value and its value as a complete object (Pickford, 2010). In the antiques trade, objects of poor quality or limited demand may be sold for their metal content alone. This is referred to by the self-explanatory term of "scrapping". Eqs.(3.1) and (3.2) formalise this. The scrap value of an item is important in ascertaining a floor price for it (Draper et al., 2018; Stephens, 2017).

$$\text{Fine Silver Mass}^{1} \approx \text{Mass of an Object} \times \text{Purity} \tag{3.1}$$

$$\text{Scrap Value} = \text{Fine Silver Mass} \times \text{Silver Price (localised currency)} \tag{3.2}$$

It can be useful rather than to consider price directly, to examine the auction/sale price as a mark-up (function) of the scrap price. One possible measure of this would be given in Eq.(3.3) which is defined over the interval $(-\infty; +\infty)$ and replicates the calculation of cost mark-up (between sale and cost price). Obviously when the premium is 0 there is no difference between the sale and scrap price. A premium mark-up of 10% would indicate that the sale price is 10% higher than the scrap price.

$$\text{Percent Premium Mark-up on Scrap} = 100 \times \frac{\text{Sale Price} - \text{Scrap Price}}{\text{Scrap Price}} \tag{3.3}$$

The data used in this Chapter comes from Stephens (2017) and unless otherwise specified refers to all origins. Table 3.1 gives the summary statistics for mass and fine mass. These two variables track in a near linear fashion because the different purities are in the region [0.800, 0.925]. Quartile 1 is of keen interest because objects below this size would be considered 'objects of virtue' by DMs, rather than silverware (which would refer to larger wrought silver). Such objects of virtue would include vinaigrettes, snuff boxes, pickle forks, konfyt[2] forks and other examples. The distinction is informative for the problem structuring because it indicates that there are actually two problems here: the pricing of objects for which the scrap value has a real influence (viz. objects of larger silverware) and the pricing of objects for which this is irrelevant (viz. objects of virtue). Section 5.2 explores the problem structuring challenges and Section 6.4 the multicriteria decision modelling difficulties that the DMs encountered.

[1] The approximation sign is used deliberately because in small objects of silver (for example cigarette cases) the entire object is made of silver; but, this relationship tends to weaken with objects of higher mass. In these larger objects we encounter some oddities. For example, the use of bone or horn is very common in large silver objects. Silver is very conductive of heat. Bone or horn is used as "stoppers" or "joiners" to connect, for instance, the handle of a teapot with its body. (This prevents the user from burning themselves.)

[2] A thick, chunky, specifically South African fruit preserve.

Table 3.1: Summary statistics for mass and fine mass

Variable	minimum	q1	median	mean	q3	maximum
Mass	1.70	164.75	378.5	831.64	948.00	12510.00
Fine Mass	1.44	148.69	342.4	750.31	868.81	11571.75

To further illustrate this problem, Figures 3.1 and 3.2 are illustrated for the domains (fine mass in grams) [0, 1250] and [150, 1250]. In Figure 3.1 the relationship is monotonically increasing between price and fine mass. If one focuses on the domain [150, 1250], a near linear relationship seems appropriate. Unlike the relationship between price and fine mass (Figure 3.1), the relationship between price and premium mark-up (Figure 3.2), is not monotonically increasing. However focusing on the domain [150, 1250], a near linear, monotonic relationship again seems appropriate. The premium mark-up changes dramatically from a mean response of around 6000% mark-up (above scrap price) at 150g to a mean response of around 200% mark-up at 1250g. The confidence interval shows that the premium becomes more volatile at larger values of fine mass. These insights highlight that preferential independence will need to be addressed.

From Figures 3.1 and 3.2, an insight relevant to problem structuring is that the use of currency value (as the dependent score) would not be appropriate. Potential alternatives could include the historically favoured currency per gram measure. However this would suffer from the same shortcomings as the mark-up premium in that it is still preferentially dependent. Hence the nature of the score required for the MCDA process is a score that clearly has a relation to currency value but is somewhat more abstract (in order to be preferentially independent). Part of the methodology (Chapter 4) in particular the problem structuring, was seeking such terms. Section 5.2 provides the results thereof.

These results are dominated by British (origin) silver. The following section contrasts and compares Cape (origin) silver to British. In doing so it highlights that certain origins do not experience the relationships demonstrated above. This fact is revisited in Section 6.5.

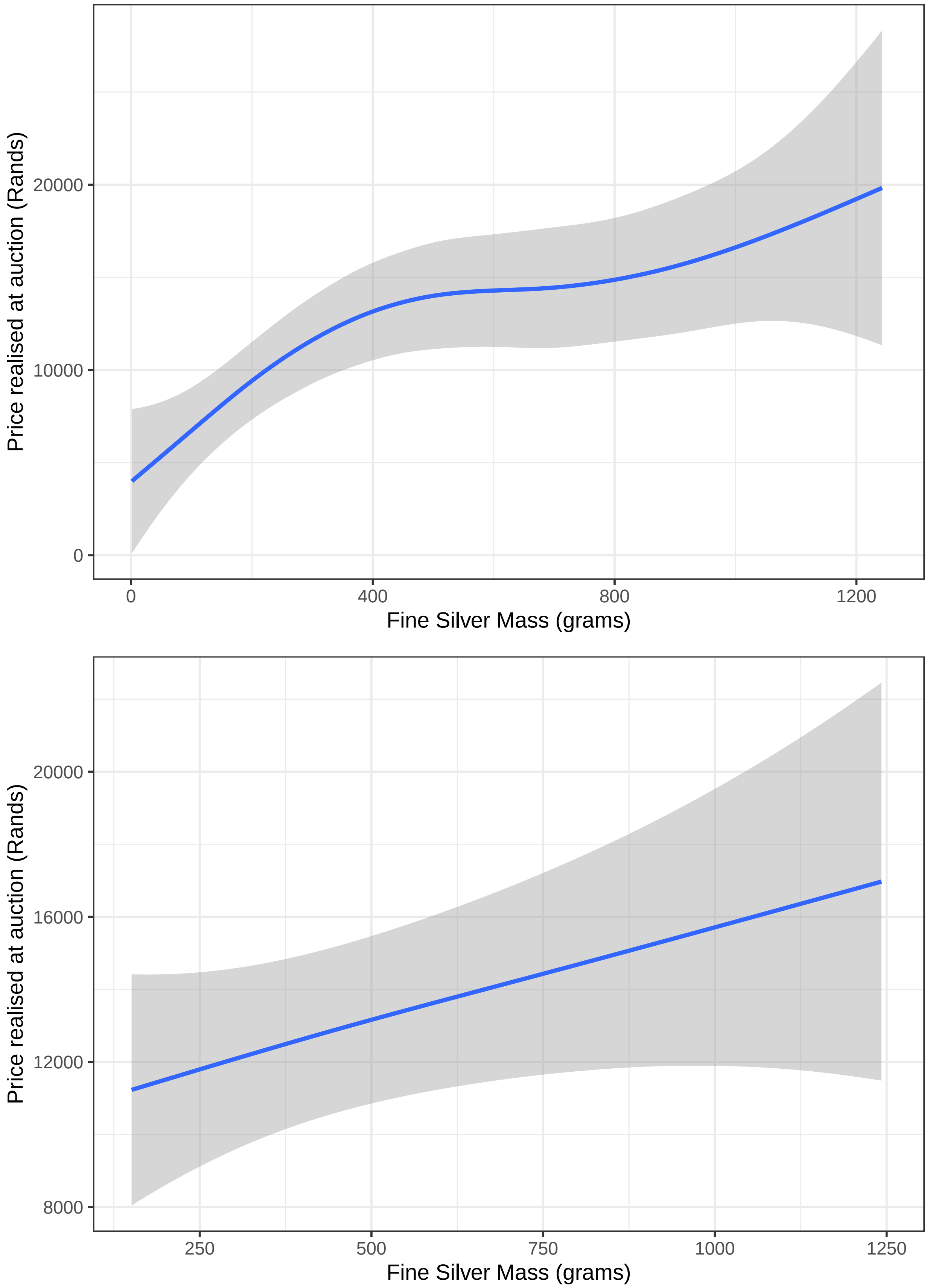
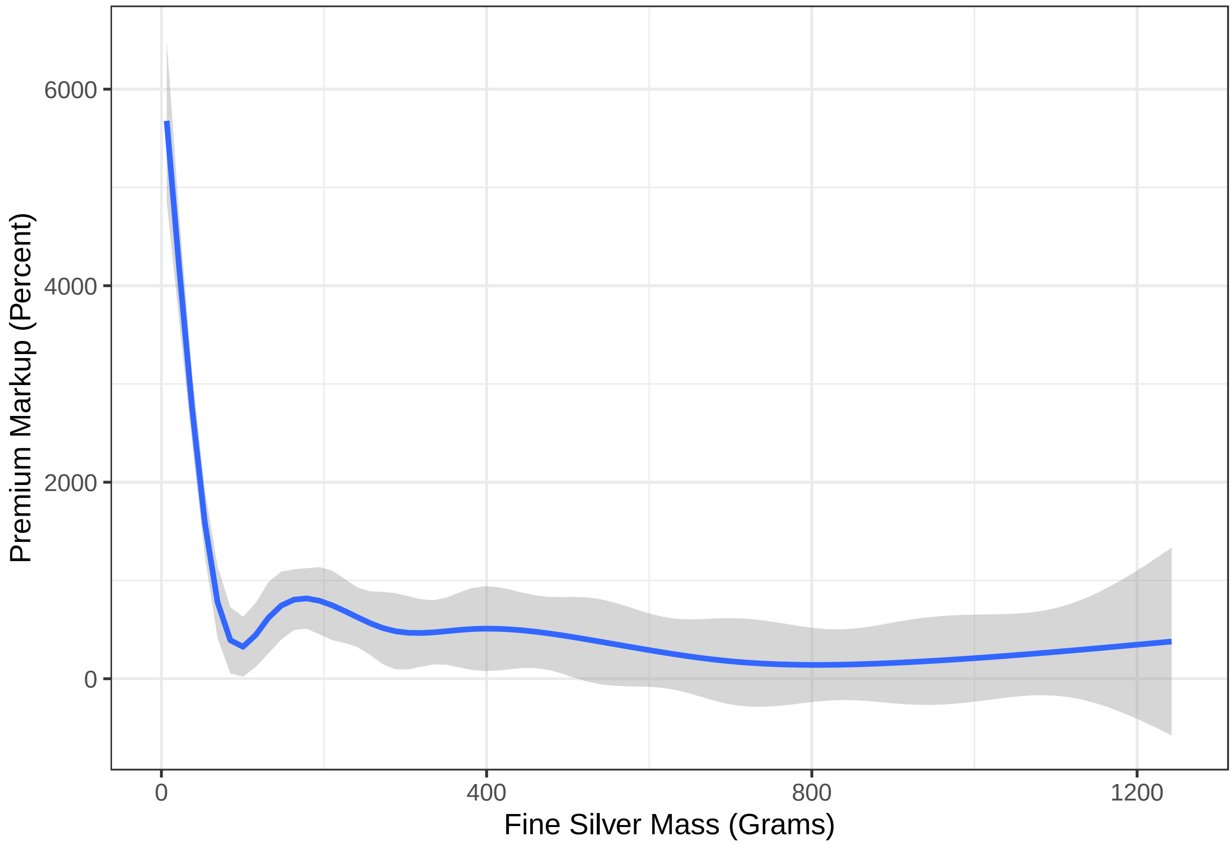


Figure 3.1: Fine Mass and Auction Price. The relationship is plotted as a smoothing spline GAM with a 95% confidence interval given by the grey region. The top figure has a domain of [0, 1250]. The bottom figure has a domain of [150, 1250].



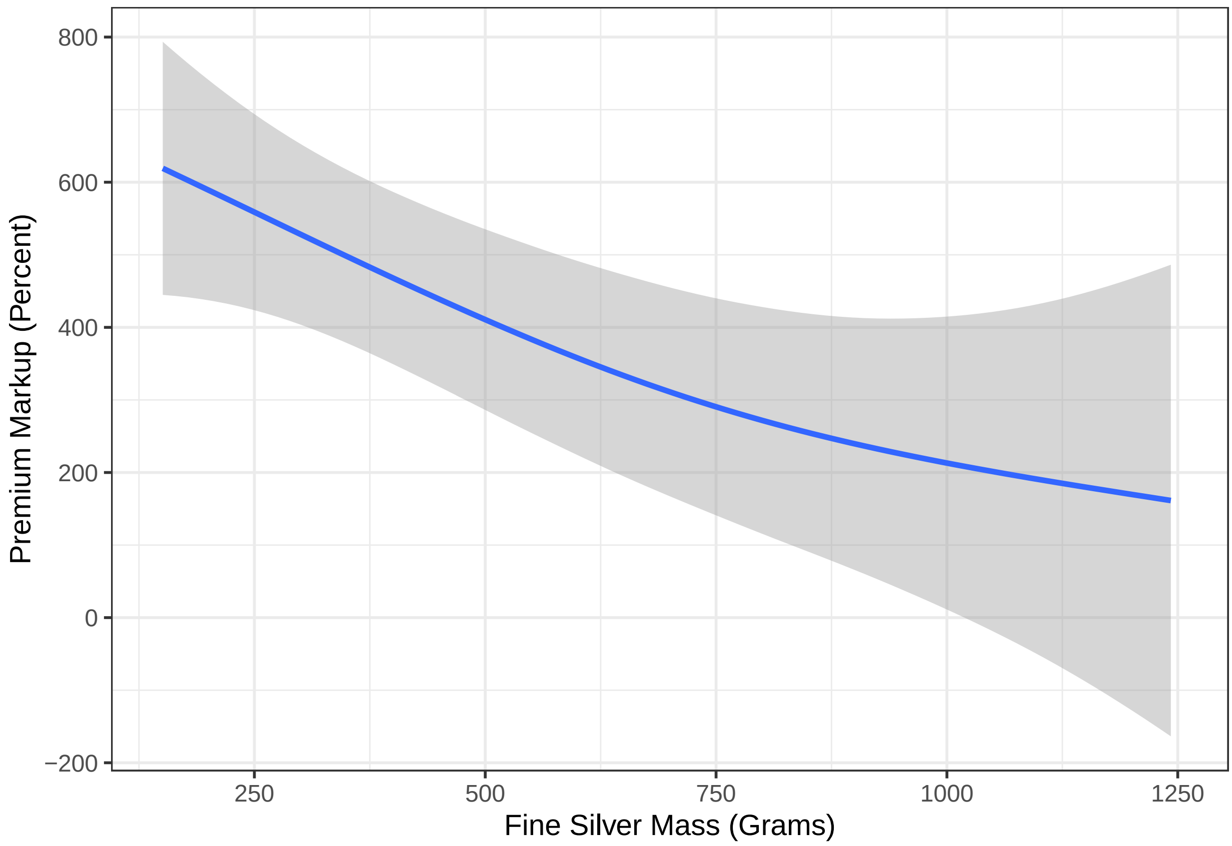


Figure 3.2: Fine Silver Mass and Premium Mark-up. The relationship is plotted as a smoothing spline GAM with a 95% confidence interval given by the grey region. The top figure has a domain of [0, 1250]. The bottom figure has a domain of [150, 1250].

3.1 Cape Data

Figure 3.3 repeats the analysis of Figures 3.1 and 3.2 for both British (origin) and Cape (origin) silver. Similar patterns to those evident in Figures 3.1 and 3.2 hold for the British (origin) silver, but Cape (origin) silver behaves differently.

Two things characterise this difference. Firstly, the functions for Cape (origin) silver are larger in magnitude than those of British (origin) silver (viz. for both auction price per gram and premium mark-up). The premium mark-up for Cape (origin) silver is considerably larger (a factor of 400 times larger at certain masses) – critically without the declining trend towards zero, evident in British (origin) silver, which suggests a decoupling between the premium mark-up and mass[3]. Secondly, the wide confidence bands present for Cape (origin) silver, in conjunction with the behaviour of these functions (unpredictable turning points and changing gradients), suggest that – unlike for the British case – one is not able to replicate this behaviour using well-known functions.

Once the Cape silver data is divided by category as is done in Table 3.2 (page 14), it is easy to see (even visually) that mean mass has a small correlation with the mean premium for any given category. Using the data in Table 3.2 a linear regression (formally Mean Premium = Mean Mass + c, with c as the regression constant) is calculated and reported in column one, Table 3.3 (page 14). The correlation coefficient is 0.09 between mean mass and mean premium (for a given category). This supports, as already argued, that the premium has become decoupled.

What about the relation between mean mass and mean price for a given category? The effect of mean mass on mean price is more strongly correlated (correlation coefficient of 0.3), but only peripherally statistically significant, as Table 3.3 column two demonstrates. Formally this linear regression is given by Mean Price = Mean Mass + d, with d as the regression constant. Once the size of objects in the different categories has been accounted for (column three, Table 3.3), the effect of mean mass is no longer statistically significant. However, the size variable is statistically significant. Formally this linear regression can be expressed as $\text{Price} = \text{Mean Mass} + \text{Size}_{\text{medium}} + \text{Size}_{\text{small}} + e$, with e as the regression constant and where $\text{Size}_{\text{medium}}$ and $\text{Size}_{\text{small}}$ are dummy variables. This highlights the fact that mass tracks an object's size (volume) and once size is taken into account the price may not be that closely related to mass. In fact, of all three regressions the one that takes the size of the objects into consideration (column 3) has the greatest explanatory power. The role that the size of objects play in the multicriteria modelling process is dealt with in more detail in Chapter 7. In review this analysis shows that the effects of mass can be understood through size, and size can be incorporated into the attribute category for Cape silver.

[3] Multicriteria decision modelling interviews will confirm that this is what the DMs expect.

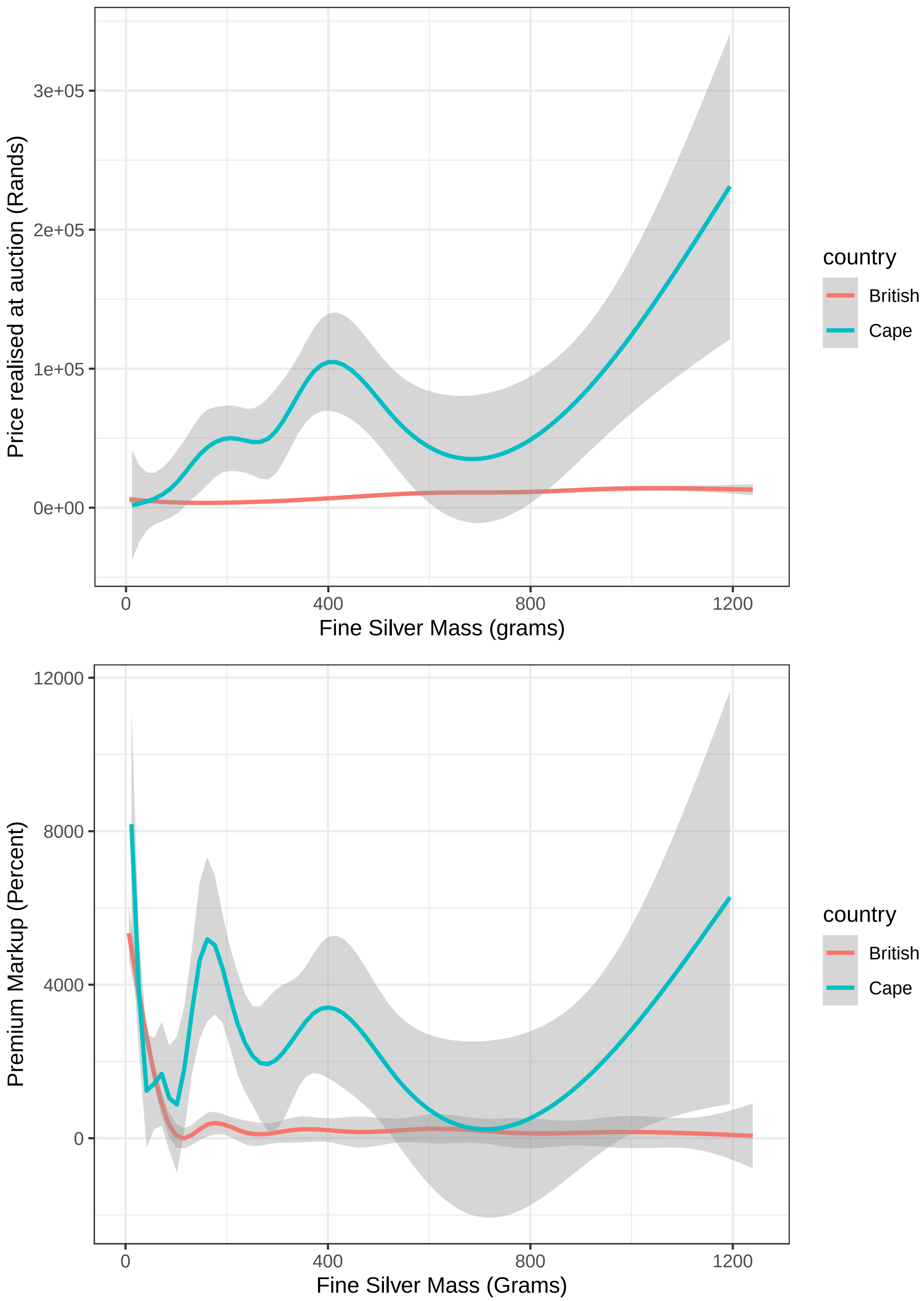


Figure 3.3: The top figure gives the function of Fine Silver Mass and Auction Price, the bottom figure gives the function of Fine Silver Mass and Premium Mark-up. The relationships are plotted as smoothing spline GAMs with a 95% confidence interval indicated by the grey region. Cape data is given in blue, while British data is red. For both figures the domain is [0, 1250].

Table 3.2: A summary of the mean premium, mean mass and mean price in Rands for different categories of Cape (origin) silver (original n = 188).

Category	Size	Mean Premium (%)	Mean Mass (g)	Mean Price (Rands)
Bible Clasps	small	660	46	8782
Dinner Services	large	6148	290	96540
Drinking Vessels	medium	3425	176	64748
Pap[1] Boats	medium	849	110	7612
Salvers & Trays	large	2264	400	62530
Small Boxes	small	2659	69	9550
Table Silver	small	535	89	7991
Tea & Coffee Wares	large	6036	371	143926
Vases	large	1125	2223	110000
Wine Labels	small	7239	15	7612

[1] A porridge usually made with maize meal.

Table 3.3: Linear Regression on Cape (origin) silver. Data source: Table 3.2

Dependent variable: (column)	**Mean Premium** (1)	**Mean Price** (2)	**Mean Price** (3)
Mean Mass (g)	0.016	42.658*	5.247
	(0.017)	(22.822)	(17.730)
Size: Medium			−63,511.840*
			(27,661.040)
Size: Small			−90,745.290***
			(24,460.050)
Constant		35,763.260*	98,941.310***
		(16,741.760)	(20,464.220)
R^2	0.088	0.304	0.790
Adjusted R^2	−0.014	0.217	0.685

Note: *p<0.1; **p<0.05; ***p<0.01

Chapter 4

Methodology

This study, over the course of 9 months, took place in Cape Town, South Africa. Identification of possible respondents occurred through a thorough search of the directory of registered antique dealers on the South African Antiques Dealers Association website (South African Antiques Dealers Association, 2019). The keyword silver was used as the preliminary criteria. This search yielded 15 potential respondents across South Africa, the majority of whom were based in Cape Town with others based in Johannesburg. Then the criteria of expert[1] was applied[2], which reduced the number of possible respondents. Next, DMs with high standing in the antique silver community were identified. All of these DMs were approached to determine their willingness to participate in the study (ensuring that no potential respondents in Cape Town were overlooked). An exhaustive list of possible respondents based in Cape Town consisted of eleven respondents. No sampling techniques were employed, rather an attempt at a census was conducted. The core group refers to eight of these DMs who were available and had completed the problem structuring before November 2019.

The methodology roadmap is: i). structure and understand the problem (Section 4.1.1); ii). construct a MCDA (an additive multiattribute value theory model) for score (Section 4.2); iii). if possible relate score into price (Section 4.2); iv). use individual insights to build some understanding of the market (these are conclusions formed in Chapters 7 and 9).

4.1 Structuring and Understanding the Pricing Problem

Structuring and understanding the pricing problem took two different probing avenues: formal problem structuring (see Subsection 4.1.1) and a detailed questionnaire for each DM. In the problem structuring interview process a questionnaire was first presented. This was intended as preparation for and augmentation of the formal problem structuring. The detailed questionnaire would typically take a DM forty minutes to complete. It began by asking the DM about his/her career and the evolution of their business. Thereafter it was divided into four major sections: World View on Value of Antiques, The Market for Antique Silver, Decision Aids for Antique Silver and Pricing Decisions.

The first questionnaire section, *World View on Value of Antiques*, investigated the DM's thoughts about the use of intuition and systematic thinking in determining a price. It also sought to understand if a DM's estimates are rigid or flexible, and probed for concepts that may be useful to explore more in the formal problem structuring. The second questionnaire section, *The Market for Antique*

[1]Defined earlier to be either retail shop owners or auctioneers whose businesses focus on antique silver i.e. antique silver typically makes up the majority proportion of their business.

[2]In many instances applying this criteria required a physical site visit because it was not always clear either from catalogues or online stock sheets.

Silver, requested background about the contemporary antique silver trade. *Decision Aids for Antique Silver* (the third section) asked questions about how DMs engage with the currently available decision support tools (these were described in Chapter 1), as well as whether changing fashions influence the use of these tools. The final questionnaire section, *Pricing Decisions*, aimed to gauge if the DM prices in a currency per gram framework i.e. a premium mark-up on the silver scrap price (see Chapter 3). As previously noted this was historically how objects of antique silver were priced. This questionnaire section also sought to understand if the DM considers himself/herself a *price-setter* or *price-taker* and against which reference point the DM's pricing decisions are formed. An interesting final question is whether the respondent sees himself/herself as adding value through provenance (this term is explained in Table 5.1, page 22). The full questionnaire can be found in Appendix A.

Together, the formal problem structuring and the questionnaire attempt to provide a fairly holistic view of the antique silver trade. In particular, they aim to understand the focal question: how the describable features of an antique silver object are used to determine a perception score (or utility) which can ultimately be expressed as (or converted into) a monetary price.

While formal problem structuring and the questionnaire provide some insight, to probe this question directly requires the use of MCDA. However, MCDA is unsuccessful unless good problem structuring occurs. Therefore the joint use of these PSMs and MCDA methods are required.

4.1.1 Methods for Problem Structuring

Many problem structuring methods (PSMs) as well as MCDA techniques are appropriate in instances where actions (alternatives) are well defined. For example, when trying to choose the location of a new global headquarters in Europe for a large corporation, the alternatives could be the set $\{London, France, Paris\}$. For the DMs, price determinations are a repeated activity which perhaps occurs several times a day, each time with a different object. Therefore, defining a set of alternatives is impossible. The situation sounds perfectly appropriate for what Keeney (1996) would have described as "value-focussed thinking". Further, the problem investigated by this work involves either DMs who own retail stores and work in isolation (typically in competition with one another) or some DMs who price for auction houses – neither involve groups of stakeholders nor DMs who buy into collective solutions.

Marttunen et al. (2017, 3), in an extensive review of MCDA and PSMs applications, gives the following as methods of problem structuring: (a) cognitive maps and group maps (Pidd (2009) would characterise this as SODA I), (b) DPSIR framework and PSR framework[3], (c) Scenario Planning[4], (d) Soft Systems Methodology, (e) Stakeholder Analysis, (f) Strategic Assumptions Surfacing and Testing, (g) Strategic Options Development and Analysis (SODA) (Pidd (2009) would characterise this as SODA II) and finally (h) SWOT Analysis.

SODA I and SODA II are both causal cognitive mapping exercises[5] whose roots lie in cognitive psychology, which is a discipline focused on understanding how people think (Pidd, 2009, 109). This family of PSM would be best suited to the problem sketch because it allows for deep and flexible engagement. The objective of SODA (both I and II) is described by Rosenhead and Mingers (2001,

[3]Best suited to problems with well defined alternatives. Fairly flexible in the application of many problematiques, but designed for interactions between society and the environment.

[4]Best suited to problems with well defined alternatives. Well suited in the application of the choice or portfolio problematiques especially with long-term strategies in identification and analysis. Importantly forces the consideration of improbable, but potentially devastating scenarios.

[5]The cognitive map consists of concepts (the graph's nodes) and pointed arrows (the graph's edges) where the arrow head points from a concept to another concept in the direction of causality. Concepts should be linked together so as to imply causal action. Concepts are bipolar in nature with the opposite pole being given after a series of dots (. . .) should it not be obvious. For example, *'client aggravated . . . client feels comfortable'*. Although very infrequent, some cognitive maps have other types of edges: straight lines or causal lines with a T. In this case the connected concepts are correlated or temporally causal respectively.

36) as a deliberate attempt to increase the problem's complexity rather than reduce it. It aims to capture the DMs thinking about a focal question so as to later develop an explicit model. Note that the map itself cannot be considered a model of thinking, rather it is thought of as a reflective tool (Pidd, 2009, 110).

SODA II is a PSM which starts at the group level and attempts to collectively create the representation of a complex system of strategies and goals in the causal framework (Marttunen et al., 2017; Pidd, 2009; Rosenhead and Mingers, 2001). In contrast, SODA I can be dissected into three parts: individual cognitive mapping, merging these cognitive maps into a 'strategic map', and a subsequent group stage with this strategic map (see Figure 6.8 in Pidd (2009, 127)).

4.2 Methods for MCDA

The main technique in this research is the application of multiattribute value theory (MAVT). The objective is that the DM build a model to seek a score (which will then be related to price). The type of model chosen here is the additive model. The form of the additive model, for a particular alternative a, is given by Eq.(4.1):

$$V(\mathbf{z}(a)) = \sum_{i=1}^{m} \tilde{v}_i(z_i(a)) \,, \tag{4.1}$$

where the measure of performance (score) is denoted by $z_i(a)$ for each criterion i[6]. The function $\tilde{v}_i(\cdot)$ weights each performance measure appropriately. The vector $\mathbf{z}(a)$ is defined as the vector of the particular alternative a's scores for each criterion $z_1(a), ..., z_m(a)$; and $V(\mathbf{z}(a))$ returns the overall resultant model score.

The choice of an additive model is motivated by the following reasons. Firstly, the additive model is relatively simplistic in form and very transparent[7], making it especially attractive to non-mathematical DMs with no experience in MCDA. Secondly, even if the problem structuring highlighted no criteria that DMs thought violated preferential independence, a further advantage of the additive model is that it is relatively robust to violations of preferential independence (Stewart, 1996, 1999) (should they arise). The additive model also supports the use of additional interaction terms which can be included if necessary.

The criteria are divided into continuous and categorical[8]. Breakpoints play a role in both. Each continuous criterion i has a vector $\mathbf{Z}_i = [z_i^0, z_i^1, ..., z_i^{K_i}]$, where K_i+1 is the number of "breakpoints" and $z_i^0 < z_i^1 < z_i^2 < z_i^3 < \ldots < z_i^{K_i} \equiv z_i^*$. The breakpoint is a specific level (domain value) of the performance measure z_i, for which a value elicitation will be sought. Linear interpolation will occur between the different breakpoints. In this way a piecewise linear value function approximation is formed. Providing breakpoints ensures that the DM considers values which other DMs found important for that criterion: this formed part of the problem structuring process described in Section 4.1.1. For a categorical criterion, the vector $\mathbf{Z}_i$ consists of the discrete set $[z_i^0, z_i^1, ..., z_i^*]$ with each value in the set being a breakpoint.

In the absence of a set of defined alternatives the criteria needed to be defined on a global scale. For convenience and ease of explanation to the DMs, the zero point is associated as the "worst" outcome (for that criterion) and 100 as the "best" outcome. The definition of the global scale typically requires more work than a local scale (Belton and Stewart, 2010, 32).

[6]For example, consider the alternative $a = a_1 =$ a konfyt spoon: for the criterion $i =$ hallmark clarity, the alternative a_1 has a value of 80 which results in a measure of performance (score) of $z_{i=\text{hallmark clarity}}(a_1) = 50$.

[7]For example the analytic hierarchy process (AHP) as Saaty (1980) describes, can be a far more difficult model to articulate.

[8]An example of a continuous criterion is *age*, and an example of a categorical criterion is *category*.

Criteria are divided into those that can be evaluated in a piecewise linear approximation (continuous criteria) and those presented as different alternatives (categorical criteria). Heeding the advice of Belton and Stewart (2010, 73) four piecewise segments would be used for all continuous criteria. Table 5.2 gives the $\mathbf{Z}_i$ vector for continuous criteria to be evaluated in piecewise linear value function approximations.

4.2.1 Indirect preference elicitation methods used

Indirect model assessment will be used. Two methods are proposed to determine the piecewise linear value function for each performance measure $z_i(a)$: MACBETH and a discrete choice experiment (see Section 5.3).

Measuring Attractiveness by a Categorical Based Evaluation TecHnique (MACBETH) is method for the assessment of additive value functions, as well as a software suite which requires only qualitative judgements about differences of attractiveness given as a pairwise comparison (Bana E Costa et al., 2012). The criteria importance weights ($\tilde{v}_i$) can also be determined using MACBETH. The pairwise comparisons are expressed as differences with qualitative levels: "no" [difference], "very weak" [difference], "weak" [difference], "moderate" [difference], "strong" [difference], "very strong" [difference], "extreme" [difference]. The direction of the difference is implied by the ordinal nature of the breakpoints. MACBETH translates these judgements into a series of interrelated linear programmes. The exact mathematical formulations can be seen in Bana E Costa et al. (2012).

The motivation for MACBETH is that the majority of DMs could be described as *mathematically uncomfortable*. None had used MCDA before. MACBETH provides a relatively soft introduction to MCDA for these DMs by allowing the facilitator to ask for non-numerical qualitative judgements. After the first value function was generated using MACBETH, time was taken to explain the concept of a value function with the aid of the newly created graphic result. Thereafter direct value methods (such as bisection methods) could be used to sketch value functions. Any completed value function generated using MACBETH was also confirmed using the bisection method in reverse. For example, by asking *was the increase in score between* z_i^0 *and* z_i^1 *indeed smaller than the increase in score* z_i^1 *and* z_i^2*?* This process also allowed the DM opportunity to finely adjust the scores from the MACBETH generated scores if the DM desired.

Chapter 5

Problem Structuring & Discrete Choice Experiments

This chapter presents the problem specific details of the PSMs (described theoretically in Section 4.1.1) as well as the results of the problem structuring. The results of the PSM are also given in the form of the breakpoint table, Section 5.2, which was previously referred to in Section 4.2. A methodology, design and analysis for discrete choice experiments (Section 5.3) closes this chapter. It is important that this section is presented as part of this chapter as it has considerable overlap with other sections here.

5.1 Problem Structuring Used

The first two stages of SODA I are employed in this work. The subsequent group stage normally serves as a refinement and validation stage, however for this problem sketch the purpose of the PSM is to develop a list of attributes for a multiattribute additive value model. This model was discussed in more depth in Section 4.2. In this instance the strategic map serves as a means of prompting DMs to consider missing attributes from their own individual maps and as a springboard for their own development prior to the introduction of the formal multiattribute additive value model.

The cognitive mapping exercise (which immediately followed the questionnaire) was interactively completed using the Decision Explorer (Banxia Software Ltd., 2019) software to increase speed and manage the emerging map's complexity. With the core group, the questionnaire and cognitive mapping exercise took place in the final stage of the first interview (which was relatively brief, typically lasting an hour and half). The focal question for the mapping exercise was: "explain how you currently price objects of antique silver". These maps are often highly personal and sometimes idiosyncratic. They then needed to be distilled into the strategic map (Figure 5.1, page 21) which is discussed shortly.

For the DMs whose interviews fell outside the time-frame of the core group these problem structuring exercises preceded the MCDA exercise in the same interview. Any additional concepts mentioned by that DM would then available to the MCDA, but not added to the already formulated strategic map.

A core part of the cognitive mapping exercise was the elicitation of aspiration and nadir levels for each of the concepts, the locations of significant change within these concepts, as well as the terms one could use to describe each of these changes. The word *breakpoints* (which will later be defined more formally) is used here to describe the divisions within a concept. An important suggestion was that different values are assigned to the breakpoints in order to give better discrimination in the MCDA model (Belton and Stewart, 2010, 75). An illustrative example for the concept 'age' could be the breakpoint terms: *modern, vintage, antique*; and these terms could be associated with certain values such as modern is between 0 and 50 years old, vintage is 50 to 100 years old and antique is 100 years and older.

5.2 Results of the Problem Structuring

A sample of the DMs individual cognitive maps is made available in the Appendix B. All the core members' maps were joined and aggregated into a consolidated strategic map (Figure 5.1). Table 5.1 gives a short description for each of the concepts.

An important insight of Chapter 3 is that the score of an object should have a relation to currency value, but is itself more abstract than currency value so as to not suffer from preferential independence problems. DMs suggested various terms to describe score, some of the more noteworthy were: "collector interest", "demand for ownership" or "ideal example". The journey to formulating these terms was more important than deciding on the best individual term. Through the process of having to explain to the DM the difficulties of preferential independence and the need to seek a term of mutual understanding between the DM and the facilitator, the DM was well placed to implicitly interact with the next stage (the MCDA process) without needing to have a watertight definition of the particular term being used.

In Figure 5.1 and Table 5.1 the concept *category classification* reflects the major categories rather than the sub-categories. *Category rarity* encompasses the fact that each category may have differences of rarities of these sub-categories, for example: Konfyt[1] forks vs ordinary tablespoons. Together, the concepts category classification and category rarity are meant to capture the overall category effect. The two attributes contribute directly to the score concept. This is revisited in Section 6.4.

Another problem structuring result was the extensive discussion of the criterion *silver mass* in determining the final price. The two proceeding chapters have both identified (in the literature and empirically) this role. The DMs explained that the role of silver mass was most prevalent with British silver, which would (especially in larger pieces) tend towards the binding lower bound (scrap value). When the DMs were asked how they would relate the silver mass to one of the noteworthy terms for score, DMs explained that it was not possible to think of "large" or "small" objects as attracting more or less "collector interest" (or any other score term). Later, in the multicriteria decision modelling process, trying to relate mass to score resulted in confusion and frustration. Section 6.4 provides more detail.

Motivated by the original difficulty of relating mass to a score, an alternative method of indirect preference elicitation using discrete choice experiments was explored. The initial principle advantage of this method is that DMs can be asked to make discrete choices based on a direct question, namely "which item would be valued more", rather than considering mass's effect on a MCDM with the intermediate score.

[1]A thick, chunky, specifically South African fruit preserve.

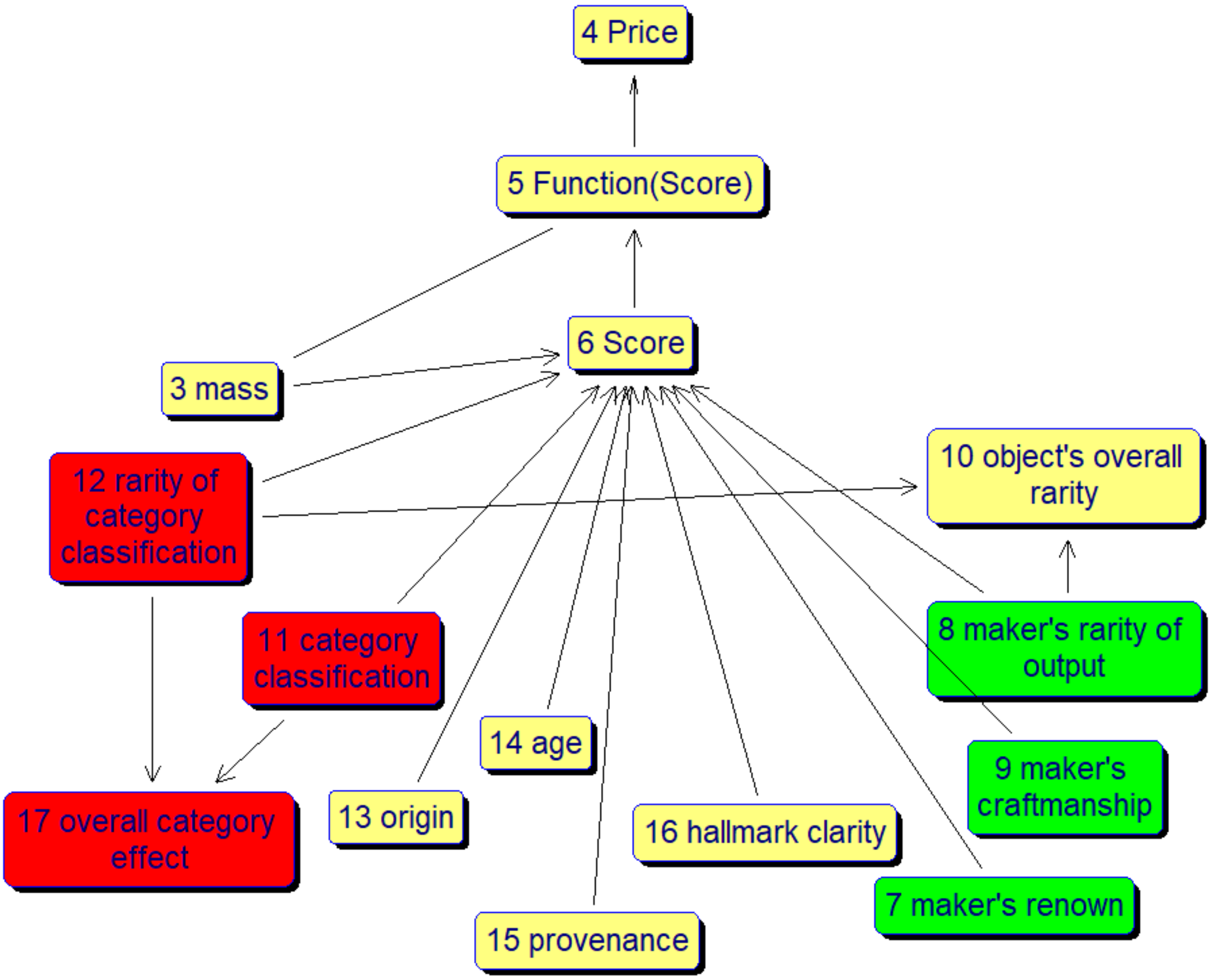


Figure 5.1: Strategic Map: Consolidated problem structuring. Concepts below the score concept are generally attributes (the inputs to the multiattribute value function model). Those attributes with a green background relate to the maker. Those with a red background to the category of the object. The score – in terms of "collector interest", "demand for ownership", "ideal example" or something similar – is determined through the multiattribute value function. Interesting to note: an *object's overall rarity* (concept 10), is what would be colloquially understood as how 'rare' an object is. This concept does not directly relate to score and is in fact understood to be the result of two rarity measures: *rarity of category classification* and *maker's rarity of output* (both of which relate to score).

Table 5.1: The possible criterion for a DM's MCDA model and their descriptions

Criterion Name	Criterion Description
age	Different collectors choose to collect certain period pieces and some periods are more popular than others (for example Georgian silver is more popular than Victorian silver).
category classification	Objects of silver can be divided into different categories and then further sub-categories. Examples of categories would be Table Silver and Teapots. Sub-categories of Table Silver would include Konfyt Forks, Marrow Scoops, Table Knives, etc.
hallmark clarity	The degree of visibility associated with the hallmark. A hallmark which is completely clear and visible (as it was on the day it was punched) is more desirable than those worn down by repeated polishing or poor initial striking.
maker's craftsmanship	The quality of silversmiths differ. More impressive workmanship as well as quality control – both concepts bound up by craftsmanship – would have originally resulted in a more expensive piece. This is still reflected in the present day in the antique market.
maker's rarity of output	The quantity of output of a silversmith varies. Some silversmiths have large teams of assistants while others work in isolation. Some collectors are interested in acquiring examples of work by all silversmiths in a particular region (e.g. Cape silver). As such, relatively scarce pieces can demand more than an identical piece from a more prodigious maker. Typically this is a concern for hallmark collectors who often collect flatware.
maker's renown	A measure of how famous/renowned a particular silversmith is. To collectors some silversmiths are well known – like Paul Storr – while others are obscure. Obscure silversmiths can create quality craftsmanship and renowned silversmiths poor quality craftsmanship. The famous female silversmith Mary Bateman, Pickford (2010) remarks as being "the least talented silversmith of the Georgian era". Nonetheless her silver remains highly collectable.
mass	Mass is an excellent proxy for an object's size. Its role in determining price is discussed in Chapter 3.
origin	Silver objects are made across the world, but some origins are of particular interest to collectors. These may be prized because of a particular style of workmanship (associated with that location) or because different origins interest different collectors.
provenance	"The history of the antique as it has moved along a train of owners, uses, locations, associations, expositions and explications from its creation to the present movement."(Rosenstein, 2009, 183)
rarity of category classification	More objects of certain categories were originally made and some categories have *resisted the melting pot* better. Different categories of silver objects have differing levels of abundance.

Table 5.2: Table with criterion breakpoints defined in the direction of increasing preferences.

Levels ($\mathbf{Z_i}$)	**zero** (z_i^0)	**one** (z_i^1)	**two** (z_i^2)	**three** (z_i^3)	**four** ($z_i^4 = z_i^*$)
age[1]	1951-present	1900-1950	1850-1899	1800-1849	Before 1800
hallmark clarity	0 clarity (indistinct)	below 50	below 70	below 90	below 100 (clear as punched)
maker's craftsmanship[2]	below 50% of makers (below average)	50% to 65% (good)	65% to 80% (meritorious)	80% to 90% (distinguished)	top 10% (exquisite)
maker's rarity of output[3]	bottom 20% (common)	20% to 40% (uncommon)	40% to 60% (rare)	60% to 80% (very rare)	top 20% (exceptional)
maker's renown[4]	bottom 50% (little known)	50% to 65% (known to specialists)	65% to 80% (recognised)	80% to 90% (famous)	top 10% (celebrated)
mass[5]	under 125g (petite)	between 126g and 275g (small)	between 267g and 600g (medium)	between 601g and 1200g (large)	larger than 1200g (extra large)
rarity of category classification[6]	bottom 20% (common)	20% to 40% (uncommon)	40% to 60% (rare)	60% to 80% (very rare)	top 20% (exceptional)

Notes:

A specified region (e.g. 50% to 65%) was easier for DMs to consider. Regions can easily be converted into statements found below the regions (e.g. known to specialists).

[1] DMs suggested an equal division between z_i^0 and z_i^*.

[2] DMs noted that at lower levels it was difficult to distinguish characteristics, while on the other hand there is much interest in the highest quality pieces and hence more discrimination was required here.

[3] DMs noted that it was easiest to think of maker's output in equal divisions between z_i^0 and z_i^*. The descriptors were debated.

[4] The rationale was very similar to footnote 2, with the need for fine discrimination for the most well known of makers.

[5] DMs suggested the statements/descriptors which define the five groups and had some intuition as to their respective cut-off points. The final cut-off points were selected using the interquartile ranges of Stephens (2017). The DMs remarked that these agreed nicely with their intuition.

[6] To reduce confusion DMs opted to follow the same construction as maker's rarity of output (footnote 2).

5.3 Discrete Choice Experiment (DCE)

A discrete choice experiment (DCE) involves the DM making a decision between two or more hypothetical choices. Figure 5.2, page 27, illustrates such a choice. The objects 1, 2 and 3, consist of criterion at specific breakpoint values. For example for object 1 "known to specialists" corresponds to z^1 for the criterion maker's renown, See Table 5.2. The DCE was to be conducted during the same session (i.e. interview) as the MACBETH modelling process. This meant it was necessary to formulate which criteria would be included prior to in depth modelling – a conundrum which would be fatally flawed. During the problem structuring process some assessment of the most important criteria were made, but this lacked the rigour of systematically determining the importance weights (done only later using MACBETH) for the additive model. Table 5.6 gives the criteria and its levels that were included in the DCE. This table is a close subset of Table 5.2.

5.3.1 Design of DCE

A chosen design for the DCE needed to be efficient (so as to maximise the certainty of the information collected with the fewest number of respondents), reflect no prior knowledge about the effect size of each z_l (Table 5.6) and minimise the total number of decision rounds (viz. discrete choices between cards) presented. The DCE experiment literature is extensive regarding how to create efficient designs where there is some prior knowledge on effect sizes (Hensher et al., 2015; Rose and Bliemer, 2009). The literature is also well developed for DCEs without priors, Burgess and Street (2005); Street and Burgess (2007); Street et al. (2005) have developed methods which are referred to as Street-Burgess constructions in this literature.

Designing a DCE in this paradigm starts by implementing an optimal linear design of generalised word lengths at least 2 (Grömping and Xu, 2014; Xu and Wu, 2001). The optimal linear design is then laid out using the Street-Burgress modulo method. A running example will be presented so that the reader can follow the method. This running example is a smaller version of the DCE calculated.

Consider three variables: maker's renown (z_1), with breakpoints [z_1^0 = Unknown, z_1^1 = Known, z_1^2 = Famous], age (z_2) with breakpoints [z_2^0 = 19th Century, z_2^1 = 18th Century, z_2^2 = 17th Century] and size (z_3) with breakpoints [z_3^0 = small, z_3^1 = medium, z_3^2 = large]. The full factorial design, which gives all possible combinations across all variables, is given by Table 5.3.

Regarding the optimal linear design, which needs to be of generalised word length at least 2, in the discrete choice literature there exists a strong preference for a balanced design (Hensher et al., 2015) [2]. The efficiency metrics of A-efficiency is defined as

$$\frac{100}{N_D} \times \frac{1}{\text{trace}((\mathbf{X^T X})^{-1})/p} \tag{5.1}$$

and the D-efficiency[3] as

$$\frac{100}{N_D} \times \frac{1}{\det((\mathbf{X^T X})^{-1})^{(1/p)}}. \tag{5.2}$$

should be optimised.

To continue the example to illustrate, a fractional factorial of generalised word length 2 is drawn from the full factorial Table 5.3. To know which rows to select from the full factorial there are two possible procedures. One would be to use a mixed integer programme with the objective of function

[2]For more details see Appendix C. In short a design is balanced when the information matrix ($\mathbf{X^T X}$) calculated from a design matrix ($\mathbf{X}$) coded with standardised orthogonal contrast coding has zero values for the diagonal elements in the intercept row and column. When a design is both simultaneously balanced and orthogonal, the $(\mathbf{X^T X})^{-1}$ matrix is diagonal and $(\mathbf{X^T X})^{-1}$ is equal to $\frac{1}{N_D}\mathbf{I_{(p\times p)}}$ (Kuhfeld, 2010, 63).

[3]D-efficiency is, in general, a relationship between $[\det(C)/\det(C_{\text{optimal}})]$ where C is the information matrix in the case of the linear model viz. $\mathbf{X^T X}$. For linear models the $\det(C_{\text{optimal}})$ is well known.

set to minimise generalised word length. Alternatively orthogonal (which are both balanced and orthogonal) arrays exist for many designs. Table 5.4 presents an orthogonal array for the full factorial in Table 5.3.

Next to convert the fractional factorial design into a linear design – hence expanding all variables into dummy variable form and applying a standardised orthogonal contrast coding – remembering that a linear design includes an intercept. The reference level for each of the dummy variable

z_1	z_2	z_3
0	0	0
1	0	0
2	0	0
0	1	0
1	1	0
2	1	0
0	2	0
1	2	0
2	2	0
0	0	1
1	0	1
2	0	1
0	1	1
1	1	1
2	1	1
0	2	1
1	2	1
2	2	1
0	0	2
1	0	2
2	0	2
0	1	2
1	1	2
2	1	2
0	2	2
1	2	2
2	2	2

Table 5.3: Table giving the different rows of an experiment involving three variables

z_1	z_2	z_3
0	0	0
2	1	0
1	2	0
2	0	1
1	1	1
0	2	1
1	0	2
0	1	2
2	2	2

Table 5.4: Table giving the different rows of an experiment involving three variables

(Intercept)	z_1^1	z_1^2	z_2^1	z_2^2	z_3^1	z_3^2
1	1.224745	-0.7071068	1.224745	-0.7071068	1.224745	-0.7071068
1	-1.224745	-0.7071068	0	1.4142136	1.224745	-0.7071068
1	0	1.4142136	-1.224745	-0.7071068	1.224745	-0.7071068
1	-1.224745	-0.7071068	1.224745	-0.7071068	0	1.4142136
1	0	1.4142136	0	1.4142136	0	1.4142136
1	1.224745	-0.7071068	-1.224745	-0.7071068	0	1.4142136
1	0	1.4142136	1.224745	-0.7071068	-1.224745	-0.7071068
1	1.224745	-0.7071068	0	1.4142136	-1.224745	-0.7071068
1	-1.224745	-0.7071068	-1.224745	-0.7071068	-1.224745	-0.7071068

Table 5.5: A linear design matrix, **X**, for the example.

expansions is z_i^0 although this breakpoint is customarily chosen any breakpoint could be used. It is this form that is denoted **X**. Table 5.5 gives Table 5.4 linear design. It is possible to confirm that Table 5.4 has a D-efficiency, Eq.(5.2), of 100% and is hence a balanced and orthogonal design.

The efficiency measures of the linear design should not be confused with the efficiency measures of the DCE – the two are independent – yet both are measures of D-optimality[4]. Depending on how a linear design is converted to a DCE even the most efficient linear design can become an inefficient DCE.

Up until now it was difficult to implement the full workflow of designing a DCE (in this paradigm) in any open source statistical programming language. The author presents the `ExpertChoice` R package which allows the user to design a DCE with ease. The package is accompanied by two vignettes: one practical, on the use of the software, and one technical, explaining (at greater length than would be permissible here) the methodology of constructing a DCE. Appendix C provides both of these vignettes and more details.

A major drawback of optimal designs is highlighted in Louviere, Islam, Wasi, Street and Burgess (2008) which demonstrates that respondent consistency declines as the design choices become more efficient. Achieving utility balance in a design, an important characteristic of optimising the D-optimality criteria of the DCE, requires that every card in a decision round has a different level for z_l. Consequently although statistically efficient the large number of simultaneously varying criteria Louviere, Islam, Wasi, Street and Burgess (2008) found to cause respondents to think heuristically. Essentially resulting in what Kahneman (2011) would describe as a "thinking fast" process rather than the intended "thinking slow" process which would have carefully weighed up the alternatives.

Street and Burgess (2007) found that an easy way to achieve utility balance was the use of a modulator pattern. Returning to our example, as each variable has three levels, the application of a modulator $(0, 1, 0)$ would be to adjust any given row in Table 5.4 by one breakpoint for the second variable. The first row would then be modified from $z_1 = 0$, $z_2 = 0$ and $z_3 = 0$ to $z_1 = 0$, $z_2 = 1$ and $z_3 = 0$. Similarly a modulator $(2, 0, 0)$ would adjust every row by two breakpoints for the first variable. The first row under this modulator would be $z_1 = 2$, $z_2 = 0$ and $z_3 = 0$. The modulator makes use of modulo operation (a mod n) where a is given in the modulator and n is the number of breakpoints in that variable. An example, if the modulator $(2, 1, 1)$ is applied to $z_1 = 0$, $z_2 = 2$ and $z_3 = 1$, the result would be $z_1 = 2$, $z_2 = 0$[5] and $z_3 = 2$. The choice of modulators which ensures complete utility balance can be shown to be the D-optimal design.

The use of the modulators $(1, 0, 1)$ and $(0, 1, 0)$ for this example would create three choice cards (the original 5.4, one modified by the first modulator and one modified by the second). The DM

[4] This is because they are based on a likelihood ratio of the determinants of achieved vs theoretically ideal design.

[5] $(2 + 1) \bmod 3 = 0$

Kindly read the descriptions of the two objects carefully. Then please answer the questions that follow.

Description	Object 1	Object 2	Object 3
Maker's Renown	known to specialists	recognised	little known
Technical Perfection (of the object) (Craftsmanship)	exquisite	below average	distinguished
Category Rarity	rare	very rare	uncommon
Size	petite: under 125g	small: between 126g and 275g	extra large: larger than 1200g
Age	early 19th Century	before 19th Century	late 19th Century

If you had to choose between these three objects which:
(please mark the appropriate box)

	Description	Object 1	Object 2	Object 3
Q1	would have the greater aesthetic response?	◉	○	○
Q2	would likely have the highest price per gram?	◉	○	○
Q3	would likely command the highest price?	○	○	◉
Q4	would you choose for your own collection if offered either one for free?	◉	○	○

Figure 5.2: An extract from the website used to capture the selection by the DMs is this example of one of the twenty questions posed to a DM in the DCE. The DM would make a selection between Object 1, Object 2 or Object 3.

would then be faced with the choice of which card to select. Based on the choices the DM makes DCE seeks to identify the partial value functions for each variable.

For the DCE conducted in this dissertation, the Table 5.6, gives the variables and their breakpoints. Table 5.7 gives the rows of the full factorial design, determined using an orthogonal array[6], used in the DCE. This is converted into choice cards using the modulators $(1,1,1,1,1)$ and $(3,3,3,3,3)$. Figure 5.2 shows the first question of twenty questions (i.e. choices) from the completed DCE design with a response. Notice that the levels between options have no overlap due the modulators used. This DCE had a D-efficiency of 98.883%. It is one of the two running examples in Appendix C practical vignette the complete calculations are given there.

[6]As such the design is orthogonal and balanced (hence D-efficiency of 100%)

z	**attribute name** (z)	z_l	**level name** (z_l)	**Description**
1	Makers Renown	0	bottom 50% of makers	common
1	Makers Renown	1	50% to 65% of makers	known to specialists
1	Makers Renown	2	65% to 80%	recognised
1	Makers Renown	3	80% to 90%	famous
1	Makers Renown	4	top 10%	celebrated
2	Technical Perfection	0	below 50% of craftsmanship	below average
2	Technical Perfection	1	50% to 65%	good
2	Technical Perfection	2	65% to 80%	meritorious
2	Technical Perfection	3	80% to 90%	distinguished
2	Technical Perfection	4	top 10%	exquisite
3	Category Rarity	0	bottom 20%	common
3	Category Rarity	1	20% to 40%	category rarity
3	Category Rarity	2	40% to 60%	rare
3	Category Rarity	3	60% to 80%	very rare
3	Category Rarity	4	top 20%	exceptional
4	Size (of object)	0	under 125g	petite
4	Size (of object)	1	between 126g and 275g	small
4	Size (of object)	2	between 276g and 600g	medium
4	Size (of object)	3	between 601g and 1200g	large
4	Size (of object)	4	exceeds 1200g	extra large
5	Age (of object)	0	1951-present	
5	Age (of object)	1	1900-1950	
5	Age (of object)	2	1851-1899	
5	Age (of object)	3	1801-1850	
5	Age (of object)	4	before 1800	

Table 5.6: Design for research on antique silver objects to be answered by experts. (A 5^5 design.)

Makers Renown	Technical Perfection	Category Rarity	Size	Age
0	0	0	0	0
4	3	2	1	0
3	1	4	2	0
2	4	1	3	0
1	2	3	4	0
2	3	4	0	1
1	1	1	1	1
0	4	3	2	1
4	2	0	3	1
3	0	2	4	1
4	1	3	0	2
3	4	0	1	2
2	2	2	2	2
1	0	4	3	2
0	3	1	4	2
1	4	2	0	3
0	2	4	1	3
4	0	1	2	3
3	3	3	3	3
2	1	0	4	3
3	2	1	0	4
2	0	3	1	4
1	3	0	2	4
0	1	2	3	4
4	4	4	4	4

Table 5.7: The fractional factorial design for the DCE. The rows are the levels z_l of the variables.

5.3.2 Analysis of DCE

Louviere, Street, Burgess, Wasi, Islam and Marley (2008) shows that efficiently designed DCEs can be used to modelled individual preferences; demonstrating this in two cases one with simulated data and another in a small test set completed for this paper. Critically though Louviere, Street, Burgess, Wasi, Islam and Marley (2008) does not compare the partial value functions obtained against those determined using another method eg. direct elicitation (bisection methods) or indirectly (eg. MACBETH). Two methods are suggested, one is the use of ordinal regression which Louviere, Street, Burgess, Wasi, Islam and Marley (2008) demonstrates using simulation yields estimates that are very close to the simulated utility functions. The use of ordinal regression for preference modelling also forms part of the multicriteria decision analysis literature – see Greco et al. (2008) for example. In fact according to Louviere, Street, Burgess, Wasi, Islam and Marley (2008) a fully D-efficient discrete choice design would yield the most tightly bound estimates possible with ordinal regression. Unfortunately, an oversight of the method for capturing the data (see Figure 5.2) is that no rank ordering was asked for from the choice alternatives. Should a rank ordering had occurred implementing the ordinal regression would have likely been the preferred method of analysis.

The alternative is the implementation of a weighted conditional logistic model; an approach based in random utility theory. Random utility theory seeks to address the fact that a discrete choice is a binary event (either a chose is taken or not taken, even when there may have been multiple options that could have been taken in any given choice set[7]) and that estimations of the choice action can be modelled. As such, models such as the multinomial logit model and conditional logistic model aim to model the action of choice probabilistically by minimising a function of the predicted probabilities for each choice given a certain choice set.

Unlabelled or generic choice data are DCE where the names of the alternatives convey only the relative order of their appearance within each survey task, (e.g., object A, Object B, Object C). With such data, as is investigated in this dissertation, a weighted conditional logistic model, Louviere, Street, Burgess, Wasi, Islam and Marley (2008) shows, is theoretically appropriate. In detail, a conditional logistic model has the same likelihood formula as a stratified Cox model with each case/control group assigned to its own stratum (the stratum being the each choice set), time set to a constant, status of chosen = case and not chosen = control, and estimation using the exact partial likelihood (Therneau, 2015; Therneau and Grambsch, 2000). The weighting is determined by counting the number of times a particular combination is picked across all the discrete choice experiments. In a DCE where the variables milk = Y, N and sugar = Y, N there are four combinations: YY, NN, NY, and YN for milk and sugar. If the combination YY is chosen ten out of twenty choice actions then its weight is 10. See Louviere, Street, Burgess, Wasi, Islam and Marley (2008, 134-140) for details.

[7]A choice set such as Figure 5.2 shows is one iteration of DCE.

Chapter 6

Phase I– Multicriteria Decision Modelling

This chapter is one of two results chapters. The first section of this chapter provides an in-depth account of the multicriteria decision modelling (MCDM) process in action, the results thereof (Sections 6.2 and 6.3 respectively), a motivation for problem restructuring and finally the restructured problem (Sections 6.4 and 6.5 respectively). Section 6.6 wraps up this chapter and is about the expectations that interactions with DMs created and how the website `CapeSilversmiths` was created to fulfil these expectations. The next iteration of MCDM took place after the problem restructuring and hence this chapter is called Phase I Multicriteria Decision Modelling.

The second chapter, Phase II, Chapter 7 takes the major insights from the revised problem structuring – that the MCDA process should occur while considering antique silver objects from a single origin – and applies this to Cape origin silver (in South Africa). A further development of that chapter is relating multicriteria model score to price.

6.1 Conducting the modelling interview – MCDM process in action

During the MCDM process, when preference elicitation was used (see Section 4.2 for the methodology), the strategic map played an important role as a review device. The strategic map gave the DM the opportunity to consider what others may have thought was important in understanding the focal question (and objective) of "how to price a piece of antique silver". Presenting the strategic map would typically take 10-15 minutes. Special emphasis was placed on showing the DM how their own concepts were aggregated to this map.

After presenting the strategic map the next stage would be for the DM to complete the discrete choice experiment (DCE). The DCE served as a good introduction to the concept of "having to make trades-offs" between the different criteria. Typically completing the twenty question DCE would take between 10-15 minutes.

Thereafter focus turned to building the additive model using MACBETH or pen and paper or a combination – whatever suited the DM best. (This is as described in the Methodology, Section 4.2.1.) This session could take anything between one hour and three hours. DMs tended to focus on concepts that they had contributed to the strategic map, but occasionally concepts that the DM did not suggest would strike a chord. Naturally, the DM always has the opportunity to apply a low or no criteria importance weight (see eq. 4.1, page 17) to any of the concepts/criteria in the strategic map which facilitated the DM taking ownership of the MCDM process. Any comments made during this process by the DM were noted. These assisted in constructing a richer understanding of the

MCDM process and form part of the results.

6.2 Results from Phase I Multicriteria Modelling

There were five DMs who interacted during the earliest of the MCDM processes. (This is as distinct from the problem structuring phase.) Of these, three are wholly completed (and presented here) and two were partially completed. The results are presented with each DM being given a different respondent identification (RESPID). The RESPIDs are constant for the whole of this chapter.

6.2.1 RESPID 3: Jacques

Jacques retails a wide variety of silver objects. His business, because of its location, has significant foot traffic. It also includes a mix of new items (as opposed to antiques) – sometimes as much as the majority of his stock. Jacques seeks fast moving stock that would appeal to "at least two types of buyer for the same item".

Table 6.1 gives Jacques' value function for the category attribute. Collector interest[1], Jacques explains, is greatest for small boxes (vinaigrettes and snuff boxes) category because these are portable[2] and have with specialist collector as well as an established literature for example Culme (2014) and Delieb (2002). Next, explains Jacques, for him is drinking vessels – a close second. Here collector interest exists because they can be functionally used, displayed and make for thoughtful and well-priced[3] gifts. Candlesticks and Centrepieces (with the former being slightly preferred) then follow in preference. Both have similar collector interest for Jacques as they can be physically displayed in similar places and both meet the desire for conspicuous ornamentation. Collector interest is lowest for Table Silver, Jacques explains because it requires constant cleaning and has little utility as: "it cannot be displayed".

Table 6.2 gives Jacques' value function for origin. Cape, Irish and British origin silver are all at or above aspiration level. Cape silver is strongly preferred over British silver with almost twice the score. French origin silver, on the other hand, does not do well in the South African market, Jacques explains, hence he avoids it. This explains its low score.

Figure 6.1 gives the value function for maker's renown, maker's craftmanship, category rarity and age. For Jacques there is less difference in score between objects from 1950 to 1900 than from 1850 to 1900. As Jacques sells objects of various age there is no reason for avoidance of younger objects (as we shall see in Clouseau, Section 6.2.2) nor a strong preference for objects of an early age. For maker's renown and maker's craftmanship the increase in score becomes larger for smaller increases of the attribute. The placing of the breakpoints, to provide better discrimination, worked well, capturing the rapid changes in score at breakpoints associated with highly desirable characteristics.

Table 6.3 gives Jacques' importance weights for each of the attributes. Age, category rarity and maker's renown were all relatively equal and make up the majority of the contribution towards overall score. Jacques explains that age is what his customers are immediately drawn to when they enter the shop – especially as similar items, one new, one old, can be viewed side by side in his store. Hence for him age is the aspect for which is the most influential on price of the object. Category rarity follows next. Regardless of what category an item is, a collector will always be interested in rare sub-categories, Jacques explained. Next is maker's renown: Collectors will always be attracted to famous makers – stocking famous makers is a good magnet for clientele, explains Jacques, and

[1]This is the hypothetical score used as the response variable. See Section 5.2 for more details.

[2]With a mass typically under 120 grams these items can be posted overseas at limited cost or moved out of the country easily he explains.

[3]Drinking vessels do not tend to be the heaviest of objects typically around 100 grams and hence have a relatively low lower bound.

collectors will ask for items by well known makers in preference to better crafted items by lessor makers.

Jacques remarked on the MCDM session that he found it challenging at times, but interesting. He was looking forward to seeing how other DMs had responded and comparing it to his own. The two most challenging aspects of the interview were modelling of maker's renown which Jacques felt was somewhat abstract and more generally in giving assessments that related to attributes where some hypothetical ordering was required (maker's renown, craftmanship and category rarity).

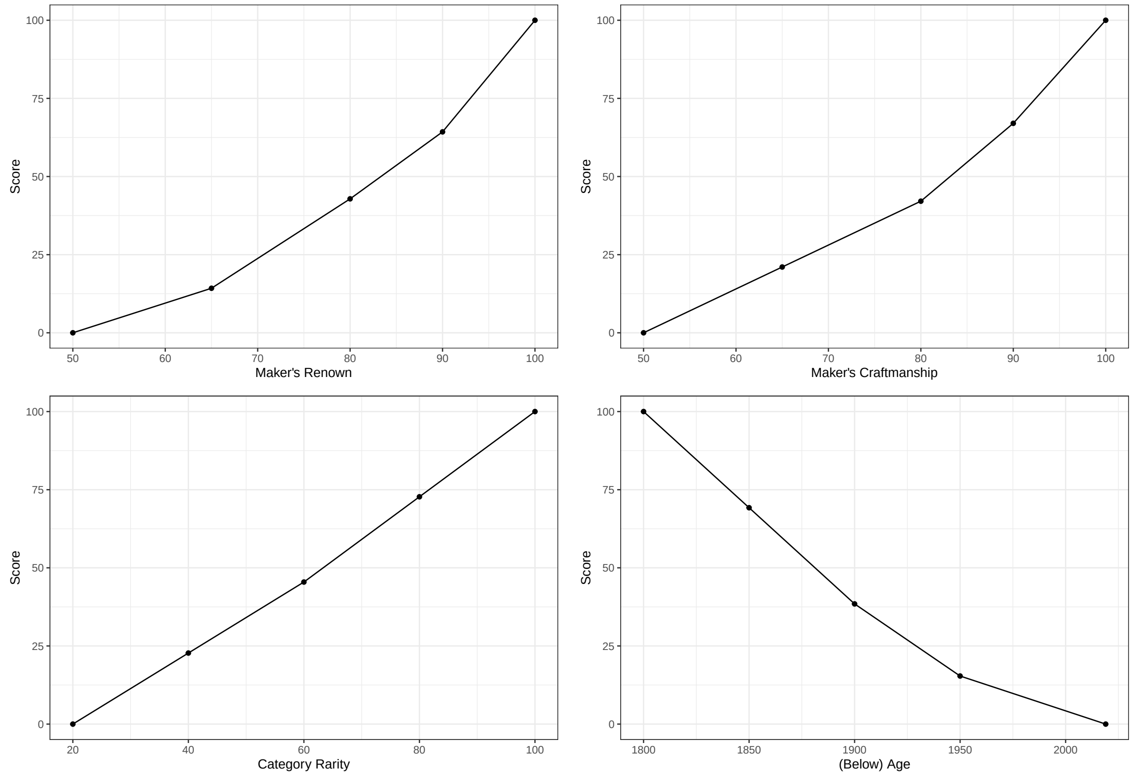


Figure 6.1: Jacques' partial value functions.

Table 6.1: Jacques' scores for each of the categories

Category	**Score**
Boxes Small	100.00
Drinking	90.62
Candlesticks	59.38
Centrepieces	43.75
Tea & Coffee	28.12
Dinner	18.75
Salvers & Trays	9.38
Table Silver	0.00

Table 6.2: Jacques' scores for each of the Origins

Origin	**Score**
Cape	217.65
Irish	158.82
British	100.00
Scandinavian	82.35
Scottish	70.42
Chinese	52.94
American	29.41
French	0.00

Table 6.3: Jacques' attributes' weights

Attribute	**Weight**
age	29.70
category rarity	25.00
maker's renown	20.29
maker's craftmanship	10.93
category	6.25
origin	4.69
provenance	3.14

6.2.2 RESPID 4: Clouseau

Clouseau's store is niche, small and targets discerning collectors. Clouseau focussed on items which were important to his business – the scores for the different category classification (Table 6.4) are very evident of this approach. For him it is clear that candlesticks, centrepieces and small boxes were of interest – other categories of silver items fell below the zero reference level given by Boxes Small. For example Table Silver and Dinner silver have large negative values which reflect the fact that for Clouseau this category of item would not ordinarily be retailed by him. In contrast to Jacques, Clouseau remarked: "Drinking vessels!? Nobody buys those". Perhaps unsurprisingly for Clouseau, category was the largest weight to his overall score.

Table 6.6 gives Clouseau's attributes' weights. For him category and age, his top two weights, are relatively equal and together account for close to 50% of the total importance weights. Category rarity, maker's renown and maker's craftmanship are all relatively equal and the next three (in decreasing order) account for close to 40%. Origin was once more one of the lowest weighted attributes.

For Clouseau, the discerning collector is most interested in objects made before 1850. Similarly with category, Clouseau would seldom retail (or purchase) items made later than 1900 – these more recent pieces fall below the 0 reference level given for objects between 1900 and 1950. A simple linear relationship in blue has been added to this partial value function, see Figure 6.2. When comparing Clouseau's value function to this line it is clear how strong his preference for earlier objects is. Objects made after 1900 received a score of 25, compared to those made before 1800 which received a score of 100.

Figure 6.2 also gives the partial value functions for maker's renown, maker's craftmanship and category rarity. Category rarity, the third largest weight hence contributor to Clouseau's overall score, has a relatively linear value function. Maker's craftmanship shows strong preference for higher scores of craftmanship especially the difference between the top 20% and top 10% of craftmanship. Clouseau explains that well crafted objects always increase the saleability of silver objects, but other attributes have more major affects on the price.

Clouseau thought that the maker's renown and maker's craftmanship were somewhat abstract concepts. He gave the example of Stuart Devlin as a maker of both maximum maker's renown and maximum craftmanship. Curiously (considering Clouseau's other value functions) Stuart Devlin was active mostly between 1950 and 1970. Clouseau went on to remark that "these are some of the only examples from that period that I purchase". When pressed for the suggestion of other makers, from other eras, Clouseau could not immediately think of another example. This led to a valuable critique from him, that the interaction be turned around and to start with the makers – grading them into the criterion's levels – then building a partial value function once there are examples of makers in each of the criterion's levels. Such a process he thought would be more "natural" of his current thinking. Clouseau described the MCDM process as at times "difficult" or "confusing". Notwithstanding this, at the finalisation of the MCDM process, Clouseau appeared very pleased with the results. "This process has helped me clarify my own thoughts" remarked Clouseau.

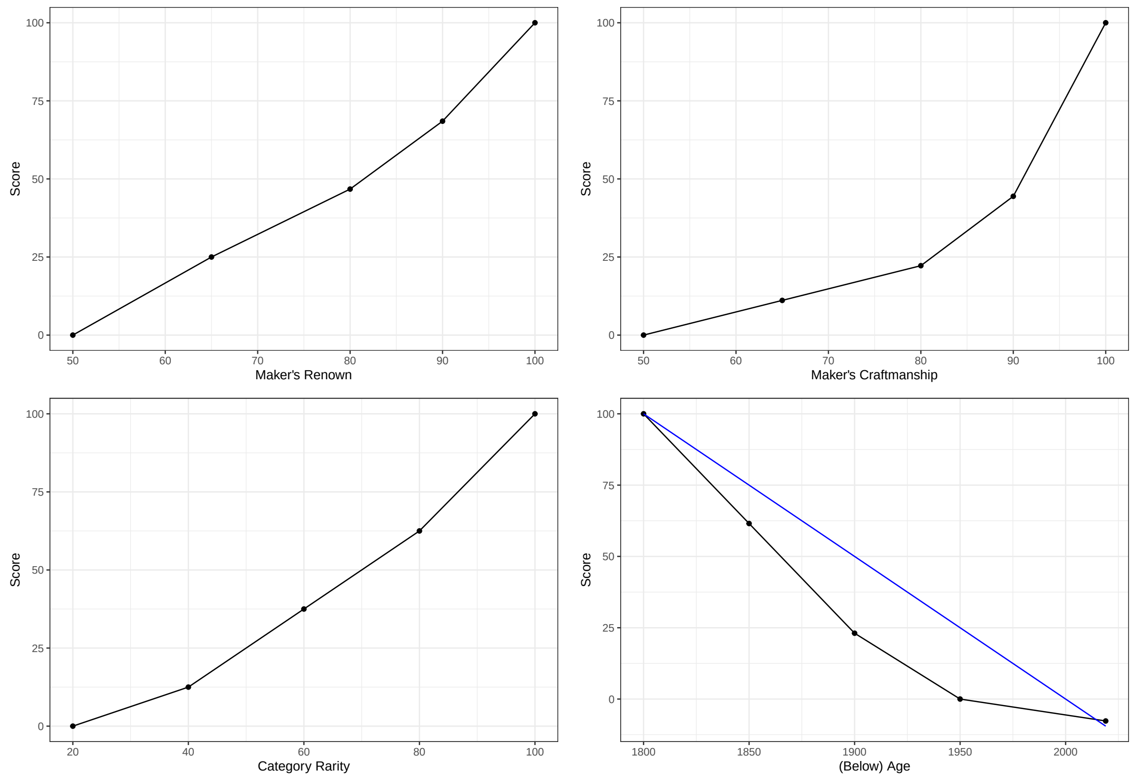


Figure 6.2: Clouseau's partial value functions. A blue linear relationship between age and score has been added to aid the discussion of this partial value function

Table 6.4: Clouseau's scores for each of the categories

Category	**Score**
Candlesticks	200.00
Centrepieces	100.00
Boxes Small	0.00
Drinking	-33.34
Tea & Coffee	-100.00
Salvers & Trays	-100.00
Table Silver	-366.66
Dinner	-366.66

Table 6.5: Clouseau's scores for each of the Origins

Origin	**Score**
Cape	257.14
Irish	185.71
Scottish	142.86
British	100.00
French	57.14
Scandinavian	42.86
America	0.00
Chinese	-14.29

Table 6.6: Clouseau's attributes' weights

Attribute	**Weight**
category	27.68
age	22.14
category rarity	15.69
maker's renown	12.91
maker's craftmanship	10.16
provenance	9.09
origin	2.33

6.2.3 RESPID 5: Hercule

Hercule has many clients including other dealers, for example, Jacques and Clouseau. This business has a high turnover of stock and Hercule needs to accurately describe each item. Hercule's sale catalogue would often be viewed before a client would see the item for sale – therefore the descriptive characteristics of the objects are relied upon to capture the reader's attention. Some clients will never view the object in person. They would be solely reliant on the catalogue and online pictures.

Table 6.9 (page 41) gives Hercule's attributes' weights. For him maker's renown, maker's craftmanship, age, provenance and category rarity cumulatively account for 90% of the weight. Each of these attributes has its weight within three percent of the next. Hercule is the first DM where provenance has a weighting above 10% of the total score.

Maker's renown, illustrated in Figure 6.3, has a similar response to the other two DMs who both reward better known makers with increasing changes in score. In explanation why maker's renown has the largest weight to his overall score, Hercule notes that it is easy to describe a maker in a catalogue: a good maker makes an impression on the reader immediately. Therefore this is something he desires and finds beneficial for his sale process.

Maker's craftmanship is Hercule's next largest weight contribution. As maker's craftmanship is defined relatively (in terms of a rank order of the pieces that the DM would originally see), for Hercule, the first two breakpoints both have a score of zero, as seen in Figure 6.3. As a means of comparison an object which Hercule may place at maker's craftmanship of 65 (or below) would certainly be in 50 (or below) for Jacques and Clouseau. The reason for this is because of Hercule's role as a sorter of the market who sees items that may enter the antiques trade, but also items that may be sold close to scrap for retail outside of the antiques trade. Hercule's maker's craftmanship score rises to a score of 25 for the 80th percentile of maker's craftmanship. By the 90th percentile for maker's craftmanship it has risen all the way to a score of 75 points. Around this level of craftmanship Hercule feels is where objects would transition into the quality demanded by the antiques trade. The increase of score from 75 to 100 occurs with a final change from the 90th percentile to the 100th percentile.

Age is Hercule's next largest weighted attribute. Once again, Hercule sees such a variety of objects that objects of different ages are to be expected. Unlike Clouseau, Hercule does not have an aversion to more modern pieces (those post 1950). Rather like Jacques, Hercule has a preference for older pieces, the changes in score are fairly linear with the changes in age, except for the change between 1950 and the present where the change in score is much less than the change between 1900 and 1950.

Hercule's business is also very different to that of Jacques and Clouseau – these differences are interesting to explore. Hercule views a lot of British silver. Unlike Jacques and Clouseau, Hercule places a much smaller score on British origin silver. (See Table 6.8 where Hercule's 100 score reference point is Scandinavian and British is 25 points above the zero score of American.) This score in part reflects Hercule's role as a sorter who would trim out pieces of British silver that trade close to or at the scrap silver price. "Whilst many items of British silver are interesting, almost as many are simply scrap" remarked Hercule. (For a more in-depth understanding see Appendix B in particular the discussion around Figure B.2.) Hercule, with his large customer base, also has interest in objects that other dealers would traditionally struggle to sell, for example, those of Chinese and Scandinavian origin. This helps explain why these origins are ranked by Hercule much higher than they are ranked by the other two DMs. Table 6.12 in Section 6.2.4 gives a more detailed comparison.

Hercule was by and large satisfied with the MCDM process. However, Hercule felt that the judgments of category classification and rarity as well as maker's renown and craftmanship needed to be locked into a particular origin for better understanding. This point is best illustrated by considering age. For Cape silver the active period is from early 18th century to the late 19th century.

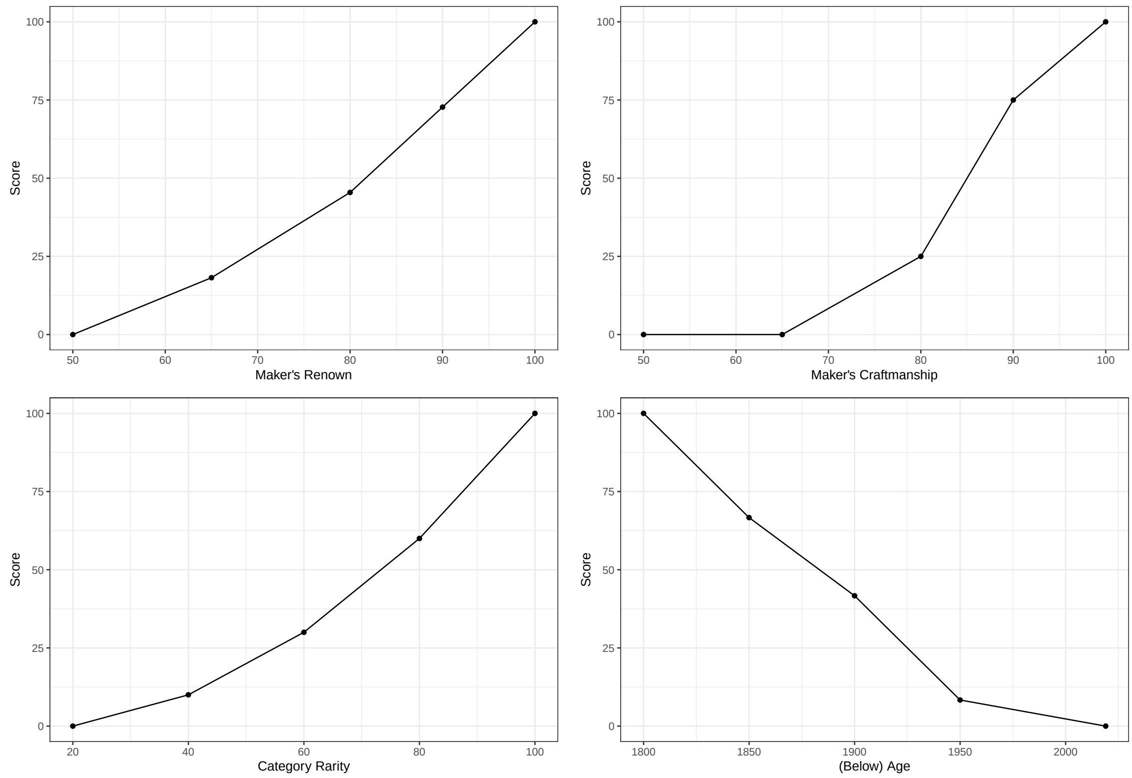


Figure 6.3: Hercule's partial value functions.

For British silver (in South Africa) the active period is from late 17th century to the present. The need to define age, irrespective of origin, reduces the needed discrimination.

Notwithstanding this, Hercule felt that these afore given partial value functions (Figure 6.3) do have "universal truth" to them despite, irrespective of the act that that they would be slightly different depending on which of the origin's Hercule was considering. This raised an interesting point, which will be further discussed in Section 6.4.

Table 6.7: Hercule's scores for each of the categories

Category	**Score**
Centrepieces	100.00
Candlesticks	90.91
Tea & Coffee	72.73
Salvers & Trays	72.73
Boxes Small	36.37
Table Silver	0.00
Dinner	-36.36
Drinking	-118.18

Table 6.8: Hercule's scores for each of the Origins

Origin	**Score**
Cape	210
Chinese	115
Scandinavian	100
French	60
Scottish	35
Irish	30
British	25
American	0

Table 6.9: Hercule's attributes' weights

Attribute	**Weight**
maker's renown	23.81
maker's craftmanship	20.63
age	17.46
provenance	14.28
category rarity	11.11
category	7.94
origin	4.77

6.2.4 Reflection on three responses: Jacques, Clouseau & Hercule

Table 6.10 compares the rankings of the attributes for Jacques, Clouseau and Hercule. (The most important attribute has a rank of 1, the least important attribute has a rank of 7.) The only attribute in the top three for all the DMs was age. Other high ranking attributes were category rarity, maker's renown and maker's craftmanship.

To return to the definition of an antique it is "a primarily handcrafted object of rarity and beauty that, by means of its associated provenance and its agedness as recognised by means of its style and material endurance, has the capacity to generate and preserve for us the image of a world now past" (Rosenstein, 2009, 260). The role of age is clear from the definition and although this definition explains why objects even post 1950 could be considered antique[4] the consistently high ranking demonstrates age's importance to this definition. In fact it is not uncommon to conflate an object's age with it being an antique. Palmer and Forsyth (2006, 236) and Stoller (1984, 26) have both given the two most common conflations, that an object must be at least one hundred years old, or prior to 1830 (Pre-Industrial Revolution), to be considered antique. These results show how deeply ingrained age is as an attribute relating to antique silver.

Tables 6.11 and 6.12 allow direct comparison of the scores of the different categories and origins respectively for the DMs. In none of these attributes did the DMs share the same or even similar rank order. Rather, these partial value functions reflect the unique DM's spheres and approaches in business. The diversity of the responses is interesting and reflects the nature of the problem derived from value. Table 6.11 demonstrates this clearly.

Table 6.10: Jacques, Clouseau and Hercule's attributes' weights. A rank of 1 represents the largest weight and 7 the smallest. Attributes arranged alphabetically

Category	**Jacques**	**Clouseau**	**Hercule**
age	1	2	3
cat-rarity	2	3	5
category	5	1	6
craftmanship	4	5	2
makers renown	3	4	1
origin	6	7	7
provenance	7	6	4

It is interesting to explore a little further. Table 6.13 gives hypothetical objects (described below) and the scores that each DM would have given. Object 1 is a piece of Table Silver, poorly made by a little known maker, typical of much churned out pre-war 1900 Table Silver. Object 2 is a

[4]For example some silver objects from the 1960s evoke the style image of that era.

Table 6.11: Comparison of category scores for Jacques, Clouseau and Hercule

Category	**Jacques**	**Clouseau**	**Hercule**
Tea & Coffee	28.12	-100.00	72.73
Table Silver	0.00	-366.66	0.00
Salvers & Trays	9.38	-100.00	72.73
Drinking	90.62	-33.34	-118.18
Dinner	18.75	-366.66	-36.36
Centrepieces	43.75	100.00	100.00
Candlesticks	59.38	200.00	90.91
Boxes Small	100.00	0.00	36.37

Table 6.12: Comparison of origin scores for Jacques, Clouseau and Hercule

Category	**Jacques**	**Clouseau**	**Hercule**
Scottish	70.42	142.86	35
Scandinavian	82.35	42.86	100
Irish	158.82	185.71	30
French	0.00	57.14	60
Chinese	52.94	-14.29	115
Cape	217.65	257.14	210
British	100.00	100.00	25
American	29.41		0

Table 6.13: Hypothetical objects and their scores from three respondents

Attributes	Object 1	Object 2	Object 3	Object 4	Object 5	Object 6
Maker's Renown	50	65	65	80	90	100
Maker's Craftmanship	50	65	65	80	90	100
Category Classification	Table Silver	Table Silver	Boxes Small	Boxes Small	Boxes Small	Boxes Small
Category Rarity	20	40	60	80	80	100
Age	1900	1850	1850	1850	1850	1800
Origin	British	British	British	British	British	Cape
Provenance	None	None	None	None	None	None
Jacques	**16.11**	**36.13**	**48.06**	**62.98**	**70.05**	**102.38**
Clouseau	**-94.05**	**-79.22**	**26.19**	**34.05**	**39.12**	**66.89**
Hercule	**8.47**	**18.27**	**23.38**	**38.37**	**55.18**	**85.91**
Mean	**-23.16**	**-8.27**	**32.54**	**45.13**	**54.78**	**85.06**
SD	**61.52**	**62.09**	**13.51**	**15.61**	**15.47**	**17.76**

better, but not excellently made, piece of Table Silver. As it is British made circa 1850 it is from the early Victorian era; a time still of extravagant dinner parties. Object 3 is similar to Object 2, but instead of being Table Silver is from the Small Boxes category with a slightly higher category rarity. This combination of category and category rarity would most likely yield a trinket box or a ladies dresser box. Object 4 has meritorious craftmanship by a recognised craftsman, once again a small box, but this time a rarer subcategory which together would mean a vinaigrette or snuff box. Object 5 is Object 4 with distinguished craftmanship by a famous craftsman. Object 6 combines the characteristics to which the DMs aspired. It is a small box (small boxes typically ranked very highly for all DMs category) of a particularly rare sub-category executed with exquisite craftsmanship by a celebrated Cape silversmith. The combination of these elements (of the attributes) would describe something like a concave snuff box. Concave snuff boxes are much more challenging to craft than ordinary (rectangular) snuff boxes.

Table 6.13 reflects how the DMs can score each of the hypothetical objects differently and sometimes radically so. Clouseau has a large aversion to Table Silver and his scores for objects that are Table Silver are negative[5]. Negative scores present an interesting conceptual area for their relation to price. In reality, Hercule, Clouseau and Jacques may if asked to value hypothetical Objects 1 and 2 agree on similar "insurance" estimates for the price. Yet Clouseau would unlikely be interested in retailing the object. Excluding the objects of Table Silver for the remaining Objects 3 to 6 the standard deviations of the score estimates remain relatively constant.

[5]A possible interpretation for a negative score would be that it models censoring in the form of items which have negative scores are not bought to retail by the DM, or that the DM would trade that item at its scrap price, rather than an incorrect interpretation which is that the DMs would place no or negative value on it.

6.3 Results from Discrete Choice Experiments

The results of the DCE are presented for Jacques (RESPID3) in Section 6.2.1. While other DMs completed DCEs their results echo Jacques and so only one DCE is illustratively reported. The methodology for the analysis of DCE is given in Section 5.3.2.

Recall that the purpose of the DCE was twofold: firstly, to test agreement of this indirect method of value elicitation (i.e. determined from analysing the choices a respondent made) against those explicitly determined (using MACBETH and other methods) and secondly, to try to gain some insight into the behaviour of mass.

Figure 6.4 gives the partial value functions from the DCE in red and, those from Section 6.2.1 in blue which were explicitly determined. For some partial value functions, such as maker's craftmanship, there is relative agreement between the DCE partial value function and the explicitly determined one. But the other partial value functions, such as maker's renown and maker's rarity of output, cast doubt about the validity of the DCE method for soliciting partial value functions with agreement to those determined using indirect value elicitation. The DCE's partial value functions exhibit unusual response patterns, including spiking points. The partial value function for the role of the object's mass has some elements of how it was expected to appear empirically as discussed in Chapter 3, but as a whole the disagreement between DCE and MCDA methods is concerning and hence suggests caution in its interpretation.

The overall DCE model for Jacques is strong with MacFadden's R^2 of 0.7712 and model p-value of 0.00281 suggesting rejection of an overall no relation null hypothesis. But the point estimates have large confidence bands and only a few are statistically significant. As the DCE already, by means of careful design, has a D-efficiency of 97.8%, improving its design further is unlikely to be a fruitful avenue of investigation.

One option would be to present the DM with more DCE choice experiments. This would occur if the original linear design (i.e. the fractional factorial) was chosen to have a word length greater than 2. However, there are already twenty questions which would take the DM between twenty-five and thirty minutes to complete. A word length larger fractional factorial would double the time requirements. Further, Kuhfeld (2010) found that DM become inconsistent and suffer from fatigue after twenty-five questions. Kuhfeld's observation also returns to the fact that should more attributes be investigated, then the number of rows (hence questions) of the linear fractional factorial required to achieve the same statistical strength would also need to increase.

In this research strong evidence of what Louviere, Islam, Wasi, Street and Burgess (2008) described was apparent in the interactive sessions. DMs would simply look for the card that had terms upon which they had decided were important without due consideration of the implications. Time taken to complete the DCE reflected this. Some DCEs would take as little as five minutes, which must be close to viable minimum for reading the information. In short it would be suggested that future designs still utilise D-efficient designs (such as the modulo method), but make use of smaller variations between options modulating one or two characteristics at a time. Having multiple options where only the differences (with respect to the first option) are highlighted may make this task cognitively easier. Furthermore this study would support the use of DCEs in experiments where there are less attributes. Use of alternative rankings within any given choice set (of a DCE) may also be a way to focus a DM into properly considering the DCE. Rankings within choice sets would also have the benefit of allowing ordinal regression estimation. The results of the DCE's determined partial value functions, presented here, are cautionary and it is suggested that more research takes place into the agreement of this method of partial value elicitation with other well-known methods.

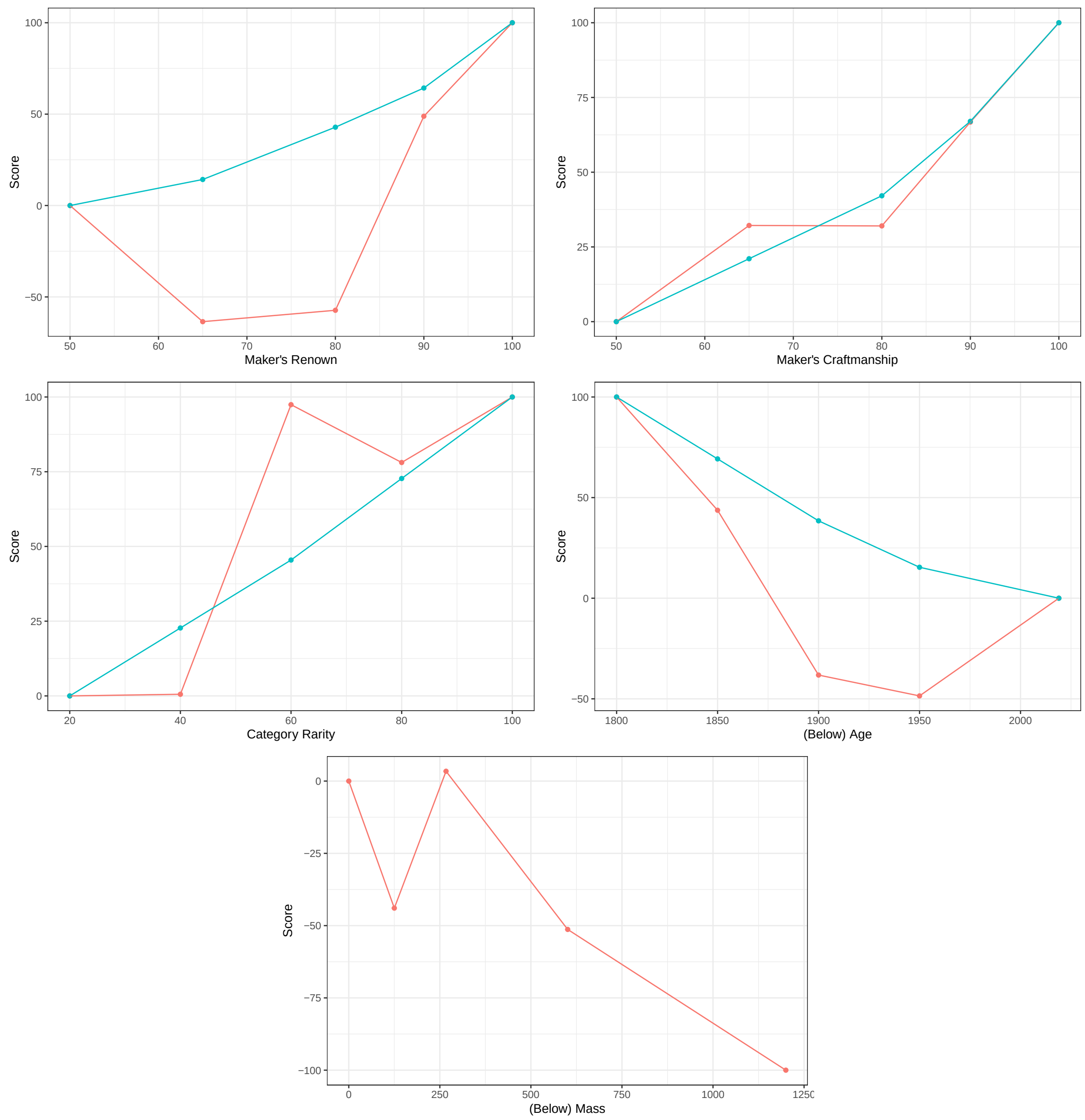


Figure 6.4: Jacques' partial value functions determined using DCE (red line) and MCDA (explicitly determined) methods (blue line).

6.4 Motivation for Problem Restructuring

Ultimately two interlinked and one independent motivation arose to suggest a restructuring. This explanation will start with maker's craftmanship, the element from which the other two elements become evident. Each origin (eg. British and Cape silver) typically has craftmanship (techniques, approach[6]) which is unique to it: within that origin differences in craftmanship quality can be identified. The comparison of maker's craftmanship between origins assumes that the characteristics of craftmanship can be globally (across all origins) defined. That is not necessarily the case. Furthermore what may constitute "good" in British (origin) silver, when considering examples of Cape (origin) silver, could be "excellent" craftmanship if achieved by a Cape maker[7]. Hence preferential independence is violated. Additionally, interactively DMs did not perceive this as a fair comparison.

Within in a particular origin the makers and quality of their craftmanship is well known to DMs. Defining maker's craftmanship within a local origin specific scale is a more realistic exercise especially for DMs who work within only one origin. The local (origin specific) scale can always be related (multiplicatively) to another origin if required. For the time being, the drawback of having a local (origin specific) scale overcomes the difficulty of comparing French to Chinese, Chinese to Scandinavian, Scandinavian to Cape craftmanship – especially, where as highlighted, each origin has specific craftmanship techniques and approaches.

After this explanation the problem is most obvious for maker's craftmanship, but also manifests itself for maker's renown. A famous maker to a DM who works in Cape (origin) silver may be an obscure maker to a DM who specialises in British (origin) silver. Even the most famous of Cape silversmiths would be unknown to a DM working in Australian silver (and vice versa) – yet both would be likely to have knowledge of British silver and, if these early MCDM sessions are representative[8], be unable to have a means of comparing them.

Maker's craftmanship and maker's renown are the two interlinked motivations. Addressing this would address Hercule's problem where the frames of reference are different for different origins. (It is surprising that the origin attribute was so lowly weighted despite how, as the DMs explained, the choice of origin (as a reference) was fundamental to how the other attributes were weighted.) The third motivation is the category attributes: category classification and rarity of category classification. It has been spread, unnaturally as Hercule and Clouseau argue, between two attributes. As the problem already requires restructuring to a single origin focus, this allows an opportunity for a revised category attribute approach. The result of this approach will be discussed in the next section. The restructured attribute should address Clouseau's problem that the earlier interactions were devoid of the actual makers and items and reliant on hypothetical ordering. This revised problem structuring avoids needing to make hypothetical orderings. Thus restructuring is an inherit modelling decision.

The MCDM process is an iterative process. Reflecting on Figure 6.5 model building, challenging thought and informing decisions is an integral part of this process. These early MCDM interviews highlighted the need, which was not apparent from the initial problem structuring that modelling (and comparisons) are most sensible within a particular origin.

A fixed origin loses inter-origin generality. The need to restructure the problem can be thought of as trading off (immediate inter-origin) generality for improved realism (for the DM) and consequently

[6]For example in British (origin) silver it is typical that the bowl of a spoon is joined to the spindle at the top of the spoon while in Cape silver the spindle is joined over the bowl. These are wholly different techniques with different pointers for well executed craftmanship.

[7]To quote Welz (1976, 32) "The work of Cape-trained silversmiths is often described as being inferior to that of immigrant silversmiths. Although there are exceptions this is generally the case. The reason for this should not be sought in the standard of training only, but rather in a combination of their training, isolation from the trends and standards in Europe and the conditions under which they lived and worked."

[8]Feedback during later interviews would indicate that they are representative.

more accurate value elicitation. To present a MCDM as a meaningful decision aid it needs to be well grounded and explanative. There is a final important motivation for fixed origin which is that it allows origin specific discrimination for attributes (such as with Cape (origin) silver where the general discrimination for the age variable was unsatisfactory).

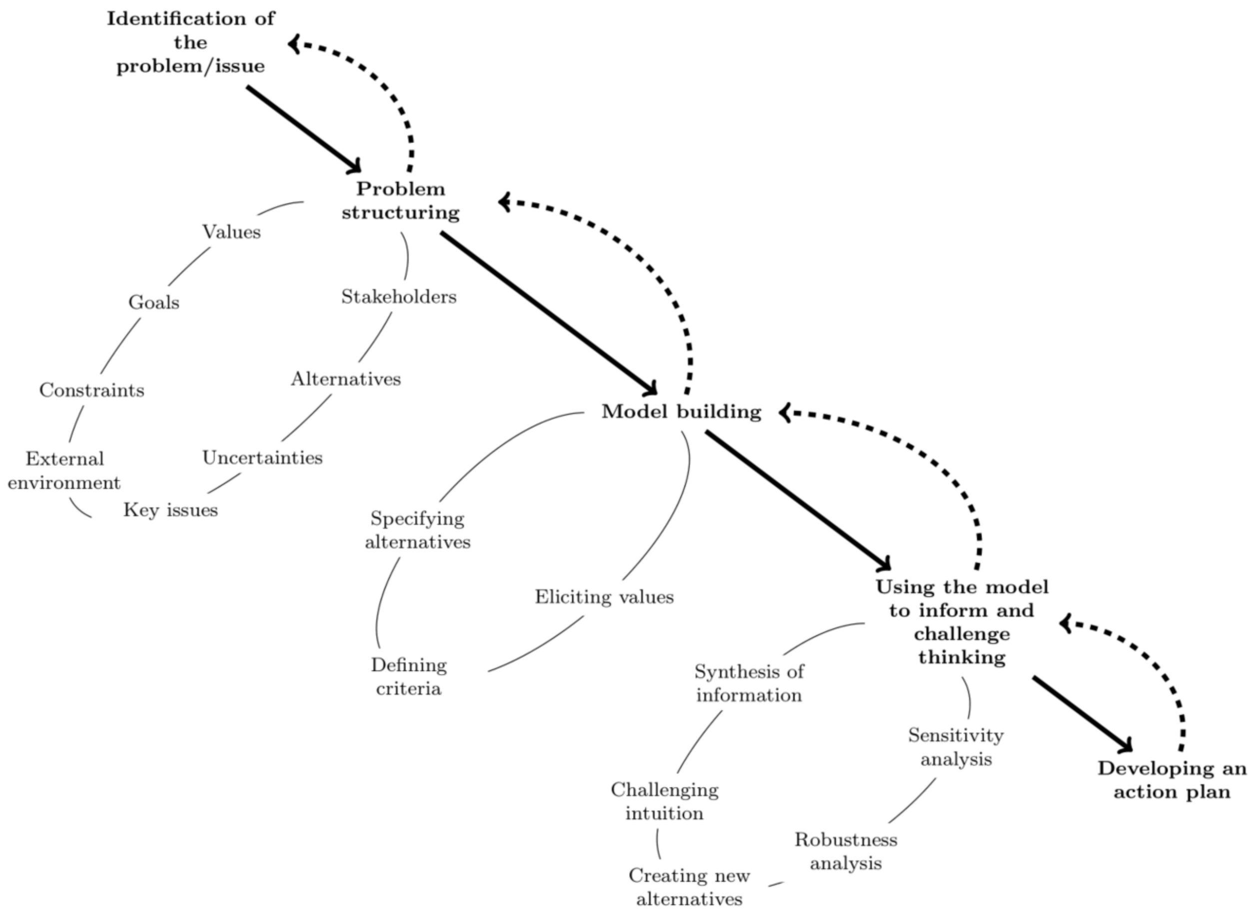


Figure 6.5: The Multicriteria Decision Modelling Process. Source: Belton and Stewart (2002, 6)

6.5 Revised Problem Structuring

The revised problem structuring incorporates all aspects discussed in Section 6.4 and the findings of the previous sections (of this chapter). Mass has been removed. As a consequence of this such a revised structuring would only be appropriate for certain origins – Where then is such a problem structuring appropriate?

The answer would be in an origin where the DM firmly believes that mass has no relation to score or the function of score (as Figure 6.6 illustrates). (See Figure 5.1, page 21, for the original problem structuring where mass is linked to both concepts.) What origin to investigate next? From the initial problem structuring (see cognitive map Figure B.1 in Appendix B) and confirmed by each of the DMs in Section 6.2 that Cape silver would be a good candidate. (More broadly, good candidates would be origins other than, in particular, British (origin) silver.)

Indeed, Section 3.1 shows beyond reasonable doubt that there is indeed no relationship between score and mass[9] for Cape (origin) silver once size has been incorporated into a measure of catagory. More generally, if DMs believe that the silver mass does not influence the price, other than as a proxy for the size of the object, i.e. it does not influence the premium function[10] (see Figure 3.3, page 13) then this is evidence that mass may be safely removed as an attribute.

A new attribute, a revised detailed category replaces the need for two category attributes. Instead of, as in Section 6.2.4, describing the vinaigrette as the composite of a category rarity classification of rare (top 20% of the distribution) and category classification of Small Box it can be directly referred to simply as a vinaigrette. The new category classification attribute will be drafted for each origin.

As Section 6.4 outlines, the revised structuring allows for all maker questions to no longer be hypothetical. Rather it shifts to assigning makers into the regions (denoted by the breakpoints) which make up the complete distribution. Within any particular region it would be possible to order the makers again, which is why a linear approximation remains appropriate – some makers would be at the bottom of a region, while others towards the top. (This process of a "second within region ordering" will not be undertaken in this study. It forms part of suggested future research, discussed later in Chapter 9.) In some origins there may be a limited number of makers, allowing all makers from that origin to be dealt with exhaustively. But in other origins if makers are drawn by simple random sampling (such as from the text Miller (2017)) then these distribution properties will still hold.

For example in the Cape (origin) silversmiths there are 55 known makers. Holmes (a DM who is the focus of Section 7.5) believes that at least 50% of the Cape (origin) silver that he sees is unmarked. As such, using the breakpoints outlined 50, 65, 80, 90, 100 as the quantiles it is easy to formulate the number of expected responses for each of the regions and consequently compare this to the expected distribution. Table 6.14 demonstrates how this distribution would look. The actual distribution after the makers have been divided can be compared to this theoretical distribution and makers can be moved around (to different regions) to satisfy the distribution. If however the DM is truly satisfied with the separation of the makers into the different regions then the breakpoints could be adjusted to reflect the DM's perceived distribution.

Figure 6.6 gives the updated and revised strategic map for an origin specific structuring and encompasses the discussion and developments of this whole chapter. The major insight of this Chapter is the origin specific problem structuring.

[9]This would also be confirmed by the DMs interviewed in Sections
refsec:Cape was to be confirmed by the DMs.

[10]A basic prerequisite is that objects of this origin consistently trade above (or are decoupled from) the scrap price.

Table 6.14: Sherlock's distribution function for maker's renown assuming that 55 makers are unknown.

Breakpoint (Quantile)	Expected Number of Makers	Cumulative Makers
50	55	55
65	13	72
80	10	88
90	6	99
100	6	110

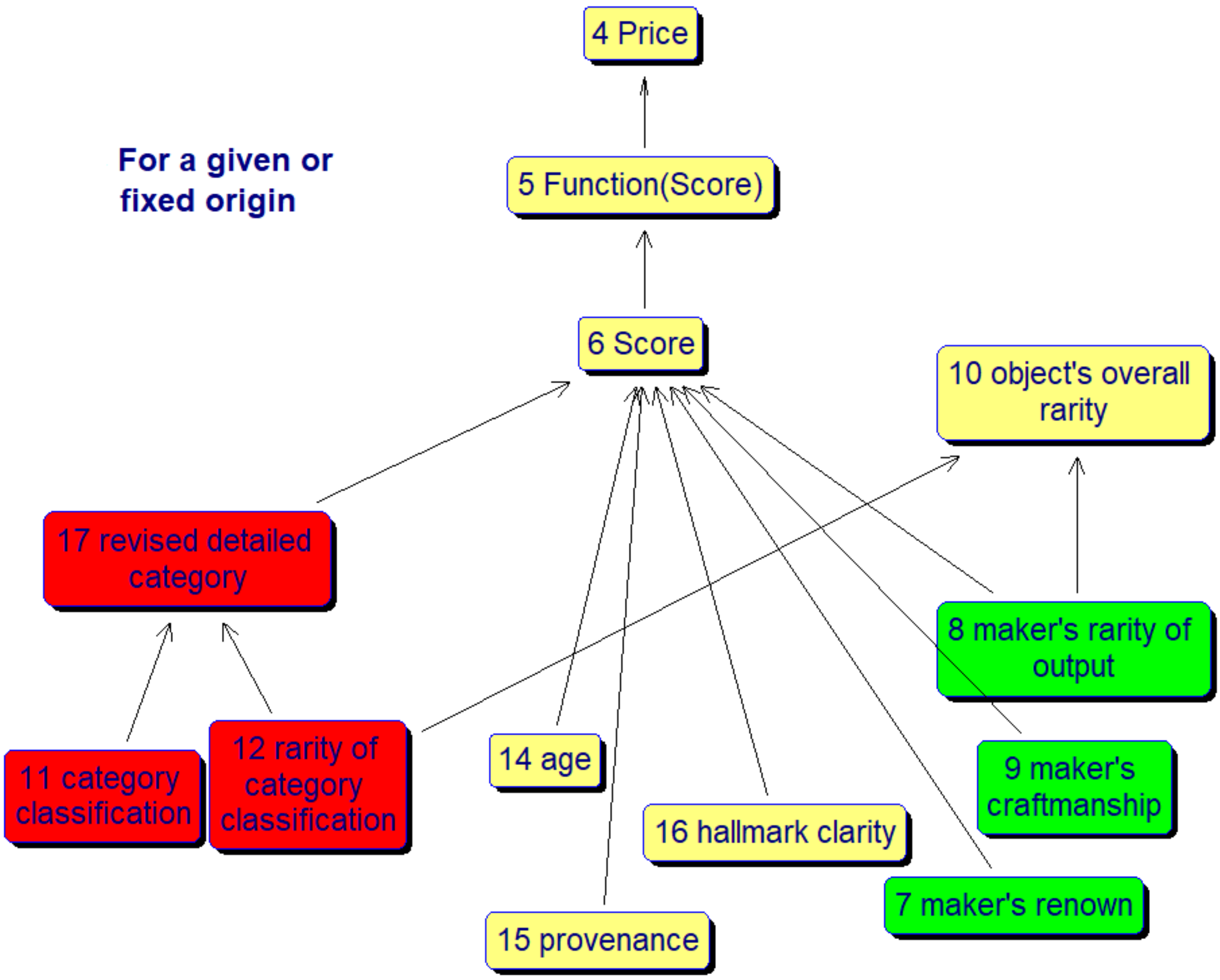


Figure 6.6: Revised Strategic Map for Origin specific multicriteria decision modelling

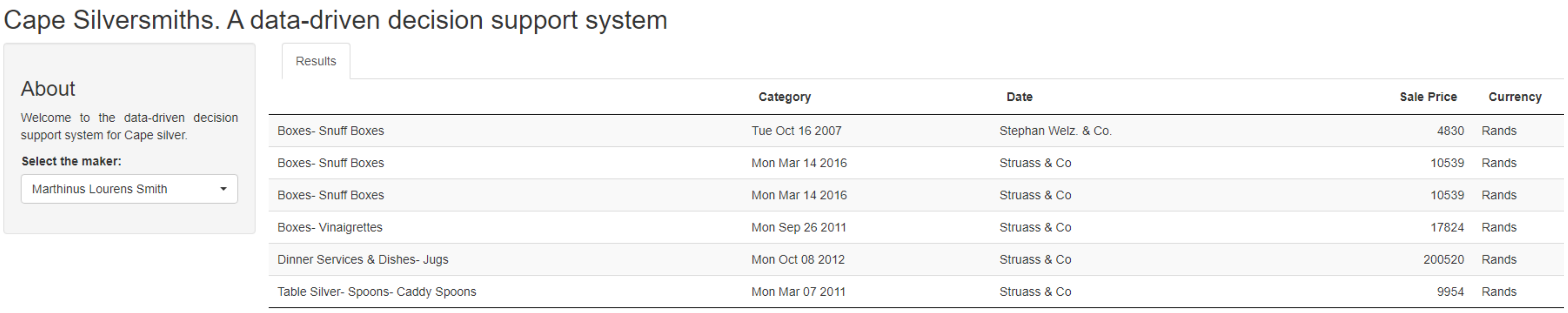
6.6 Interactions and Expectations – `CapeSilversmiths`

The antique silver retail community in Cape Town is, as already discussed, both small and well connected. The different niches within which each DM operates is highlighted in Section 6.2.4. Hercule's response (Section 6.2.3) shows a single DM who interacts with others. Hercule's position is not unique. A consistent narrative from the interviews is the DMs will purchase items from one another in the normal course of business, either through antique fairs or from their stores. Holmes (Section 7.5) explains that he competes with other DMs (that were interviewed) for the supply of goods (i.e. to purchase his retail stock from the general public).

Perhaps unsurprisingly, when working in such a tight-knit community the nature of such an in-depth study spread like wildfire. This plays out in various different ways. For example the support of certain key members was one of the reasons I could secure interviews (with others). At face value, as these interviews were unsolicited, it was not enough to secure support simply by explaining that the MCDM process may result in an interesting decision aid. (There was inherent uncertainty as to whether MCDM could work in this context and the jury was out.) Such a process has risk, (stemming from the inherit uncertainty) and the upfront time investment was not necessarily justifiable in the DM's mind. The way that I won this support (of these key members) was on the basis that the assistance was to be mutual.

Finding the appropriate reciprocation was challenging. It needed to be relevant, achievable (while this study was ongoing) and have some long-term value. Through consultation with DMs as well input from the questionnaire, in particular the section originally formulated to understand more about existing decision aids, it became apparent that there was a deficiency of historic information for Cape silver.

The decision support system, `CapeSilversmiths`, http://www.capesilversmiths.co.za was built to provide DMs with historical records from South African auctioneers for Cape silver. Figure 6.7 gives a view of a simple query being performed on `CapeSilversmiths`. Additional information is available as summary statistics. From the DM's perspective this website also increases the international profile of Cape silver, which as in Section 7.5, Holmes will remark how such a processes can affect the price of silver through the score. `CapeSilversmiths` was an interesting outcome of this research and has been appreciated by the DMs.


Cape Silversmiths. A data-driven decision support system

About

Welcome to the data-driven decision support system for Cape silver.

Select the maker:

Marthinus Lourens Smith

Results

	Category	Date	Sale Price	Currency
Boxes- Snuff Boxes	Tue Oct 16 2007	Stephan Welz. & Co.	4830	Rands
Boxes- Snuff Boxes	Mon Mar 14 2016	Struass & Co	10539	Rands
Boxes- Snuff Boxes	Mon Mar 14 2016	Struass & Co	10539	Rands
Boxes- Vinaigrettes	Mon Sep 26 2011	Struass & Co	17824	Rands
Dinner Services & Dishes- Jugs	Mon Oct 08 2012	Struass & Co	200520	Rands
Table Silver- Spoons- Caddy Spoons	Mon Mar 07 2011	Struass & Co	9954	Rands

Figure 6.7: A simple query from the `CapeSilversmiths` DSS. Additional information is available on the website such as summary statistics.

Chapter 7

Phase II– Multicriteria Decision Modelling: Cape Silver in South Africa

This chapter focusses on a single origin: Cape silver in South Africa. Few respondents would be sufficiently knowledgeable to undertake these highly focused studies; because it requires in-depth knowledge about the origin's silversmiths as well as experience handling pieces made by these silversmiths. Time availability of these respondents was often restrictive. Their knowledge is internationally sought on these subjects and, as such, the potential respondents travel frequently. Therefore this required these interviews to be planned months in advance. Notwithstanding this, the interviews were much faster to conduct than those in Chapter 6 – the extreme comfort of the DM with the subject, and the revised problem structuring lending itself to a more natural interview.

The interview process was a variation of the more general process described in Section 6.1. First, the revised problem structuring map (Section 6.5, Figure 6.6, page 50) was presented (some of the DMs had input into the earlier problem structuring process) and each DM then refined the map until satisfied. Next, refining of the attribute breakpoints occurred – the starting point for these discussions being Table 7.1 which is discussed shortly. No discrete choice experiments were conducted. Therefore upon complete satisfaction with the problem structuring, the focus shifted to discuss the makers – the next part of the interview.

The DMs were asked to rate the maker on the three (independent) characteristics: renown, craftmanship and rarity of output. The breakpoints for all the maker questions (see Table 7.1 for Cape silver) were kept constant between DMs so that different DMs could be compared. The questions were asked as a linked series (first renown, then craftmanship and finally, rarity of output) prompted by a card. Examples of these cards appear in Figure 7.1.

Check cards were also included in the pack: these were either identical or had some small variation such as including a middle name of a silversmith. See Figure 7.1 row 4 and row 3 respectively for examples. Inevitably with check cards some respondents noted that they had already answered for that silversmith. This could cause annoyance. If a DM did notice a check card an answer was asked for nonetheless.

The check cards provided some insight into a DM's answer consistency. The largest difference given was one breakpoint level. Row 3 in Figure 7.1 shows such an example. For this particular DM the DM holds the now commonly held belief that a maker described by Welz (1976) only with his initials HNS is in fact Daniel Heinrich Schmidt[1]. Despite the card Daniel Heinrich Schmidt and

[1]Daniel Heinrich Schmidt is widely considered the Cape's finest silversmith. This is significant because in this field it would be the equivalent of discovering that Rembrandt had two signatures and previous unattributed (yet

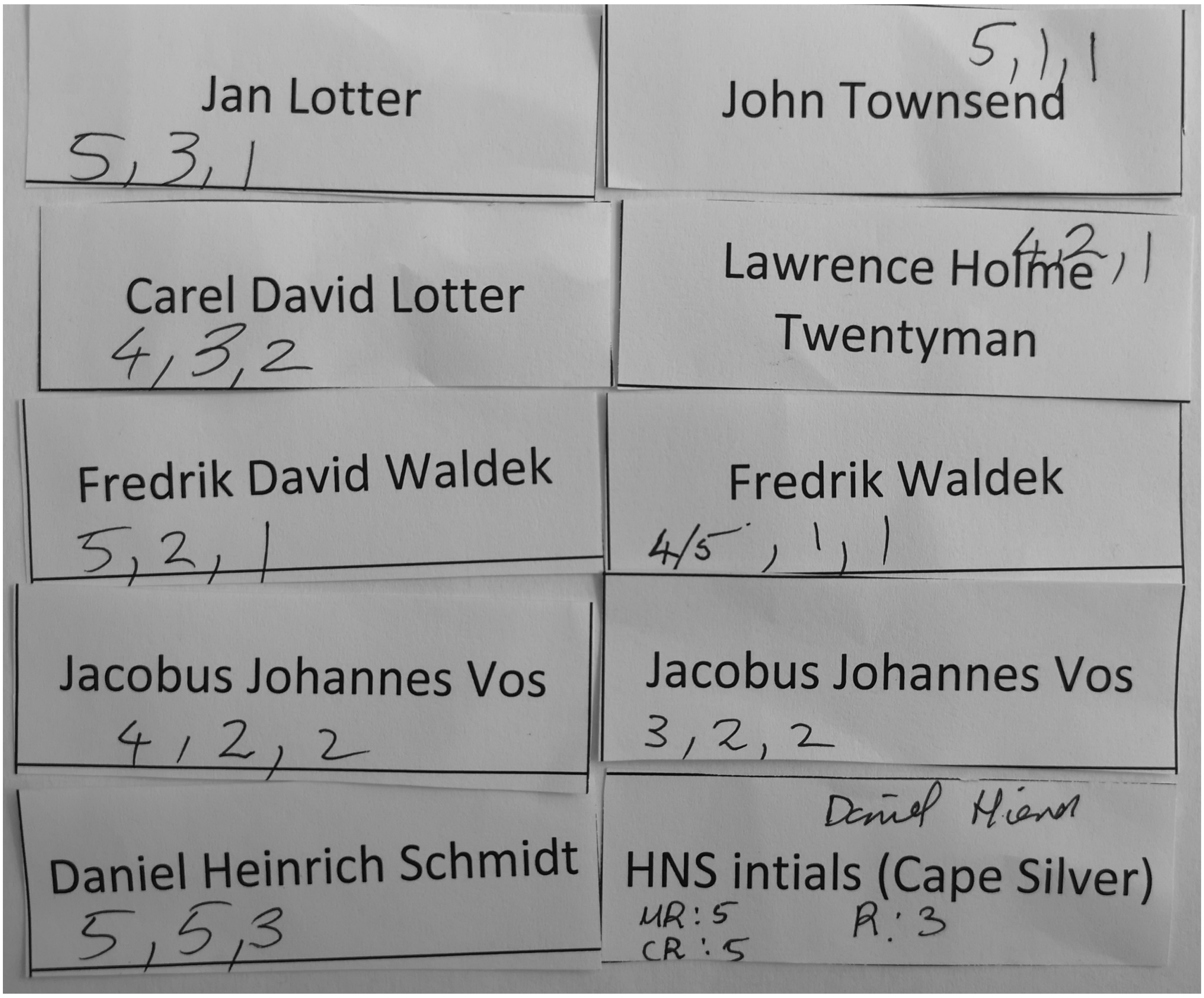


Figure 7.1: Responses to the series of questions on the makers. Each response is separated by a comma and denotes a rating given to questions of the 'renown of the maker', 'craftmanship of the maker' and 'rarity of the output of the maker'.

the card HNS initials (Cape Silver) being several cards apart the DM not only recognised them as one and the same, but gave both the same level ratings. While this example is not statistically significant it is illustrative of the degree of consistency that experts can achieve with this exercise.

Cape silver was first introduced in Chapter 2. These results are important because Cape silver is a specialist subject, with few experts and scarce historical data. Section 6.6 explains how the little historical data that does exist became a decision aid. But besides `CapeSilversmiths`, little decision support exists. The successful application of multicriteria decision aid would be advantageous for other practitioners (such as British auctioneers) who are required to price articles of Cape silver. Two DMs were interviewed, Sherlock (RESPID1) and Holmes (RESPID2). Regrettably a third potential DM was unable to make an interview after the fourth rescheduling. There are no other commercial experts in this specialised area of expertise known to the author.

To adjust to the revised problem structuring and to provide good discrimination for the Cape (origin), Table 7.1 (page 56) is a modified Table 5.2 (page 23). In Table 7.1 the age attribute reflects this much finer discrimination spanning a period from before 1700 to 1882; rather than, as originally, from before 1800 to present (Table 5.2). Table 7.1 was only suggestive and not prescriptive. DMs

finely executed) works are in fact Rembrandts.

could modify any of these, except for the three maker attributes whose breakpoints were fixed for comparison purposes.

Table 7.1: Table with criterion breakpoints for Cape silver defined in the direction of increasing preferences.

Levels ($\mathbf{Z_i}$)	**zero** (z_i^0)	**one** (z_i^1)	**two** (z_i^2)	**three** (z_i^3)	**four** ($z_i^4 = z_i^*$)
age	1820[1]-1882[2]	1801-1820	1751-1800	1701-1750	Before 1700
hallmark clarity	0 clarity (indistinct)	below 50	below 70	below 90	below 100 (clear as punched)
maker's craftsmanship	below 50% of makers (below average)	50% to 65% (good)	65% to 80% (meritorious)	80% to 90% (distinguished)	top 10% (exquisite)
maker's rarity of output	bottom 20% (common)	20% to 40% (uncommon)	40% to 60% (rare)	60% to 80% (very rare)	top 20% (exceptional)
maker's renown	bottom 50% (little known)	50% to 65% (known to specialists)	65% to 80% (recognised)	80% to 90% (famous)	top 10% (celebrated)
rarity of category classification	bottom 20% (common)	20% to 40% (uncommon)	40% to 60% (rare)	60% to 80% (very rare)	top 20% (exceptional)

Notes:
A specified percentage region (e.g. '65% to 80%') was easier for DMs to consider. Regions can easily be converted into a series of statements found below (e.g. 'meritorious').
[1] Reflects the arrival of the 1820 Settlers from Britain: known to have brought new silversmiths and changed local taste.
[2] The cessation of Cape (origin) production – See Section 2.1.

7.1 Cape Makers and their Ratings

Changes in the makers mark, style and quality of craftmanship are all contributing factors which can allow a person to differentiate between makers. Silversmiths at the Cape cannot be uniquely identified by name alone because names were sometimes common between father and son. An example hereof is Johannes Casparus Lotter (baptised 1737) and his son Johannes Casparus Lotter (baptised 1768). In order to uniquely identify each silversmith the tables that follow have surname, name and date of birth or baptism provided. Those dates which have a dagger ("†") suffix indicate baptism; as opposed to dates without this suffix which indicate birth. In instances where spelling of a name or surname is ambiguous all variants are listed separated by commas.

The Tables 7.2, 7.3 and 7.4 give the chosen breakpoint level for the attributes: maker's renown, maker's craftmanship and maker's rarity of output respectively for the two DMs in this section. These tables follow now in this section.

Table 7.2: Ratings of the (makers) renown for Cape silversmiths from various DMs

Silversmith			Makers Renown	
Surname	First Name	Birth	RESPID1	RESPID2
Ackerman, Ackermann	Christiaan	1699		
Ahlers, Alders	Oltman	1779	2	4
Bam	Johannes Andries	1749†		
Beck	Lodewyk Willem Christiaan		4	3
Beets	Daniel		3	4
Beyleveld	Johannes Hendric	1792†	1	3
Bünning	Johan Anton			4
Collinet	Daniel		2	3
Combrink	Johannes	1781†	3	4
Daniel	Peter Clark		2	
De Jongh	J		1	1
Du Moulin	Dominique Baudouin		2	4
Ficker	David			
Gaugain	Philip John			
Hasse	Johann			
Hausenius	Georg Friedrich			
Heegers	Johannes Jacobus	1778	2	3
Herman, Hermann	Frederik Lambertus	1778†	1	1
Hillegers	Frans	1744		
Hockly	Daniel	1787	3	4
Ince	Joseph		2	4
Keet	Marthinus	1790†		
Kohl	C			
Kruger	Christiaan	1761	1	3
Lotter	Carl David	1789†	3	4
Lotter	Gerhardus	1764†	3	3

Table 7.2: Ratings of the (makers) renown for Cape silversmiths from various DMs *(continued)*

Silversmith			Makers Renown	
Surname	First Name	Birth	RESPID1	RESPID2
Lotter	Jan		4	
Lotter	Johannes Casparus	1737†	3	3
Lotter	Johannes Casparus	1768†	3	3
Lotter	Johannes Marthinus	1798†	2	4
Lotter	Matthias			
Lotter	Widow W.G.	1760	1	3
Lotter	Willem Godfried	1748†		
Lotter	Willem Godfried	1780†		
Moore	William		4	
Niestad	Carel Dawid	1797†		
Schlosser	August Christoffel			
Schmidt	Daniel Heinrich	1741	4	4
Schmitzdorff	Godfried Fredrik	1777†		
Smith, Smidt, Smit	Marthinus Lourens	1722	3	3
Stephenson, Stevenson	Thomas		1	2
Steyn	Jan Coenraad			
Townsend	John		4	3
Townsend	Thomas Lock		2	4
Twentyman	Lawrence Holme	1793	3	4
Vos	Jacobus Johannes	1808	2	3
Vos	Jacobus Johannes	1834	2	3
Vos	Johan Hendrik	1749†	3	
Vos	Johan Hendrik	1806†		4
Vos	Michiel Christiaan	1759		
Vos	Widow Johan Hendrick (1749)			
Waldek	Fredrik David		3	3
Willcox	John Syms	1837		
Wolhuter	Andries Jacobus (junior)	1854†		
Wolhuter	Georg Egbertus	1792†	0	2

Table 7.3: Ratings of the craftmanship for Cape silversmiths from various respondents

Silversmith			Makers Craftmanship	
Surname	First Name	Birth	RESPID1	RESPID2
Ackerman, Ackermann	Christiaan	1699		
Ahlers, Alders	Oltman	1779	1	4
Bam	Johannes Andries	1749†		
Beck	Lodewyk Willem Christiaan		0	3
Beets	Daniel		0	4
Beyleveld	Johannes Hendric	1792†	0	2
Bünning	Johan Anton			4
Collinet	Daniel		1	3
Combrink	Johannes	1781†	2	1
Daniel	Peter Clark		1	
De Jongh	J			4
Du Moulin	Dominique Baudouin		3	3
Ficker	David			
Gaugain	Philip John			
Hasse	Johann			
Hausenius	Georg Friedrich			
Heegers	Johannes Jacobus	1778	1	4
Herman, Hermann	Frederik Lambertus	1778†		2
Hillegers	Frans	1744		
Hockly	Daniel	1787	3	2
Ince	Joseph		1	4
Keet	Marthinus	1790†		
Kohl	C			
Kruger	Christiaan	1761	3	3
Lotter	Carl David	1789†	2	2
Lotter	Gerhardus	1764†	2	3
Lotter	Jan		2	
Lotter	Johannes Casparus	1737†	2	1
Lotter	Johannes Casparus	1768†	2	1
Lotter	Johannes Marthinus	1798†	2	3
Lotter	Matthias			
Lotter	Widow W.G.	1760		4
Lotter	Willem Godfried	1748†		
Lotter	Willem Godfried	1780†		
Moore	William		0	
Niestad	Carel Dawid	1797†		

Table 7.3: Ratings of the craftmanship for Cape silversmiths from various respondents *(continued)*

Silversmith			Makers Craftmanship	
Surname	First Name	Birth	RESPID1	RESPID2
Schlosser	August Christoffel			
Schmidt	Daniel Heinrich	1741	4	3
Schmitzdorff	Godfried Fredrik	1777†		
Smith, Smidt, Smit	Marthinus Lourens	1722	2	2
Stephenson, Stevenson	Thomas		0	2
Steyn	Jan Coenraad			
Townsend	John		0	2
Townsend	Thomas Lock		1	2
Twentyman	Lawrence Holme	1793	1	2
Vos	Jacobus Johannes	1808	1	3
Vos	Jacobus Johannes	1834	1	3
Vos	Johan Hendrik	1749†	2	
Vos	Johan Hendrik	1806†		2
Vos	Michiel Christiaan	1759		
Vos	Widow Johan Hendrick (1749)			
Waldek	Fredrik David		0	1
Willcox	John Syms	1837		
Wolhuter	Andries Jacobus (junior)	1854†		
Wolhuter	Georg Egbertus	1792†		1

Table 7.4: Ratings of the rarity of the output for Cape silversmiths from various respondents

Silversmith			Rarity of Output	
Surname	First Name	Birth	RESPID1	RESPID2
Ackerman, Ackermann	Christiaan	1699		
Ahlers, Alders	Oltman	1779	3	4
Bam	Johannes Andries	1749†		
Beck	Lodewyk Willem Christiaan		0	3
Beets	Daniel		2	4
Beyleveld	Johannes Hendric	1792†	2	2
Bünning	Johan Anton			4
Collinet	Daniel		2	3
Combrink	Johannes	1781†	0	1
Daniel	Peter Clark		2	
De Jongh	J		4	4
Du Moulin	Dominique Baudouin		2	3
Ficker	David			
Gaugain	Philip John			
Hasse	Johann			
Hausenius	Georg Friedrich			
Heegers	Johannes Jacobus	1778	2	4
Herman, Hermann	Frederik Lambertus	1778†	3	2
Hillegers	Frans	1744		
Hockly	Daniel	1787	3	2
Ince	Joseph		3	4
Keet	Marthinus	1790†		
Kohl	C			
Kruger	Christiaan	1761	3	3
Lotter	Carl David	1789†	1	2
Lotter	Gerhardus	1764†	0	3
Lotter	Jan		0	
Lotter	Johannes Casparus	1737†	1	1
Lotter	Johannes Casparus	1768†	1	1
Lotter	Johannes Marthinus	1798†	1	3
Lotter	Matthias			
Lotter	Widow W.G.	1760	3	4
Lotter	Willem Godfried	1748†		
Lotter	Willem Godfried	1780†		
Moore	William		0	
Niestad	Carel Dawid	1797†		

Table 7.4: Ratings of the rarity of the output for Cape silversmiths from various respondents *(continued)*

Silversmith			Rarity of Output	
Surname	First Name	Birth	RESPID1	RESPID2
Schlosser	August Christoffel			
Schmidt	Daniel Heinrich	1741	2	3
Schmitzdorff	Godfried Fredrik	1777†		
Smith, Smidt, Smit	Marthinus Lourens	1722	1	2
Stephenson, Stevenson	Thomas		2	2
Steyn	Jan Coenraad			
Townsend	John		0	2
Townsend	Thomas Lock		2	2
Twentyman	Lawrence Holme	1793	0	2
Vos	Jacobus Johannes	1808	1	3
Vos	Jacobus Johannes	1834	1	3
Vos	Johan Hendrik	1749†	2	
Vos	Johan Hendrik	1806†		2
Vos	Michiel Christiaan	1759		
Vos	Widow Johan Hendrick (1749)			
Waldek	Fredrik David		0	1
Willcox	John Syms	1837		
Wolhuter	Andries Jacobus (junior)	1854†		
Wolhuter	Georg Egbertus	1792†	4	1

7.1.1 Comparing Judgments on Cape Makers

Table 7.5 (page 64), at the end of this section, compares ratings for each attribute between the two DMs. It does so by giving the absolute value of the difference in breakpoints. This highlights that between DMs agreement should not be taken for granted: each multicriteria model, and more finely each attribute of a model, is intended to reflect the individual expression of each DM. The maker's renown attribute had the smallest variation (SD = 0.819), this was followed by the maker's rarity of output attribute (SD = 1.14), while the most variable attribute was maker's craftmanship (SD = 1.11).

For each of these attributes the following will occur. First the correlation coefficient between the two DMs ratings will be calculated. This is to gauge if both DMs agree on the same estimate: a large correlation coefficient would indicate agreement, while a small or zero coefficient suggests disagreement. A non-parametric paired Wilcox sign test is conducted to see if there is a true mean difference (between DMs) i.e. one DM may have consistently scored all makers higher or lower. (The pairing is possible because these are repeated observations (value judgements) on the same maker.)

For maker's renown the correlation coefficient between the DMs is 0.502. Disagreement occurs most often for instances where Holmes has awarded high values for maker's renown, but Sherlock's ratings are lower. The paired Wilcox sign test returns a p-value of 0.0003512, which strongly suggests rejecting the null hypothesis – that being that the true difference in means is equal to zero. The observed difference in means is 0.85 with Holmes reporting higher values.

For maker's craftmanship the correlation coefficient between DMs is -0.077 suggesting, comparative to maker's renown, rather little agreement. The paired Wilcox sign test returns a p-value of 0.0027724, which suggests rejecting the null hypothesis – that being that the true difference in means is equal to zero. The observed difference in means is 1.1 with Holmes reporting higher values. Perhaps this is unsurprising, because there is no set criteria for the quality of a maker's craftmanship. In contrast, the renown of a maker is more easily objectified. Designing more detailed criteria for the quality of craftmanship is a point that is revisited under future research in the conclusion, Chapter 9.

Finally, for maker's rarity of output, the correlation coefficient is 0.3695, suggesting a much stronger agreement than maker's craftmanship. Although challenging because of the lack of historical records for Cape silver; maker's rarity of output is the most objective of the three maker's characteristics – since the rarity of output depends on the physical amount of output. Assuming that all Cape silver objects could be equally likely to survive until the present it would be possible to form an idea of the relative output of each silversmith by inspecting historical sale records. This is discussed in the future research section of the conclusion, Chapter 9. The paired Wilcox sign test a p-value of 0.0022331, which again suggests rejecting the null hypothesis – that being that the true difference in means is equal to zero. The observed difference in means is 0.86 with Holmes reporting higher values.

In general, the comparison between the DMs of the three maker's attributes paints the following picture: these two DMs are talking about the same market (deduced from the high agreement of maker's renown and maker's rarity of output), but with different views of the details – a level of subjectivity is used by both DMs in rating the maker's craftmanship, and Holmes systematically favours higher judgements. Part of the reason (as discussed before) is because Holmes believes the known and named Cape silversmiths are already a breakpoint separated from the unmarked and unknown marks that appear on Cape silver (See Section 6.5, page 49). The effect on the market of how the DMs rate the different makers is a subtle point. Although Sherlock and Holmes differ these factors need to be considered relative to maker's contribution to the overall price (i.e. the weight of that attribute) and subject to how the two DMs perceive the other attributes which describe an object.

Table 7.5: The absolute value of the difference between ratings from the DMs Sherlock and Holmes for each Cape Silversmith.

Silversmith			Absolute Value of Differnce		
Surname	First Name	Birth	Maker's Renown	Maker's Craftmanship	Maker's Output Rarity
Ackerman, Ackermann	Christiaan	1699			
Ahlers, Alders	Oltman	1779	2.0	3.0	2
Bam	Johannes Andries	1749†			
Beck	Lodewyk Willem Christiaan		1.0	3.0	2
Beets	Daniel		1.0	4.0	0
Beyleveld	Johannes Hendric	1792†	2.0	2.0	0
Bünning	Johan Anton		3.0	3.0	2
Collinet	Daniel		1.0	2.0	0
Combrink	Johannes	1781†	1.0	1.0	4
Daniel	Peter Clark				
De Jongh	J		0.0		3
Du Moulin	Dominique Baudouin		2.0	0.0	1
Ficker	David				
Gaugain	Philip John				
Hasse	Johann				
Hausenius	Georg Friedrich				
Heegers	Johannes Jacobus	1778	1.0	3.0	1
Herman, Hermann	Frederik Lambertus	1778†	0.0		3
Hillegers	Frans	1744			
Hockly	Daniel	1787	1.0	1.0	1
Ince	Joseph		2.0	3.0	2
Keet	Marthinus	1790†			
Kohl	C				
Kruger	Christiaan	1761	2.0	0.0	2
Lotter	Carl David	1789†	1.0	0.0	2
Lotter	Gerhardus	1764†	0.0	1.0	3
Lotter	Jan				
Lotter	Johannes Casparus	1737†	0.0	1.0	1
Lotter	Johannes Casparus	1768†	0.0	1.0	1
Lotter	Johannes Marthinus	1798†	2.0	1.0	3
Lotter	Matthias				
Lotter	Widow W.G.	1760	2.0		3
Lotter	Willem Godfried	1748†			
Lotter	Willem Godfried	1780†			

Moore	William				
Niestad	Carel Dawid	1797†			
Schlosser	August Christoffel				
Schmidt	Daniel Heinrich	1741	0.0	1.0	2
Schmitzdorff	Godfried Fredrik	1777†			
Smith, Smidt, Smit	Marthinus Lourens	1722	0.0	0.0	1
Stephenson, Stevenson	Thomas		1.0	2.0	0
Steyn	Jan Coenraad				
Townsend	John		1.0	2.0	3
Townsend	Thomas Lock		2.0	1.0	1
Twentyman	Lawrence Holme	1793	1.0	1.0	4
Vos	Jacobus Johannes	1808	0.5	2.0	1
Vos	Jacobus Johannes	1834	0.5	2.0	1
Vos	Johan Hendrik	1749†			
Vos	Johan Hendrik	1806†			
Vos	Michiel Christiaan	1759			
Vos	Widow Johan Hendrick (1749)				
Waldek	Fredrik David		0.3	0.5	2
Willcox	John Syms	1837			
Wolhuter	Andries Jacobus (junior)	1854†			
Wolhuter	Georg Egbertus	1792†	2.0		3

7.2 RESPID 1: Sherlock

Sherlock's business, like Hercule's (Section 6.2.3), is based on text and visual descriptions. But Sherlock's innovative business has an unconventional retail method for silver objects, enabling him to serve both a local and international customer base. Sherlock was drawn originally to this business because of "[a] silver object's ability to describe its past", by which Sherlock meant that the hallmarks (see Chapter 2) are expressive of where, and by whom, an object was made. The variety of categories and age of silver objects were also features which shifted Sherlock away from other collecting interests. Sherlock's business reflects many of these initial interests: it focusses on small objects and objects of virtue with clear hallmarks. In the size classification Table 3.2 (Section 3.1, page 14), it is apparent that Sherlock sells objects no larger than medium in size.

Sherlock explained that his business has three groups of customers: (1) collectors of colonial silver hallmarks, (2) collectors of Cape silver[2] and (3) family connections (to the silversmiths)[3]. Table Silver[4] is particularly appealing to the hallmark collector (group 1) – a type of collector whose interest in silver is collecting a range of exotic and rare hallmarks rather than being principally concerned with the type of object itself – because they are normally the most affordable examples of objects with a particular hallmark.

Sherlock believes that his price estimates are determined more systematically than unsystematically. He also believes that estimating price is an art, and if asked to price the same object the next day without recollection of the price he gave it the previous day, Sherlock expects there would be some variation. For objects that he personally likes Sherlock seeks a price premium which he himself does not always think is justified. Sherlock explains that he would never consider the mass (of an object) as featuring in a pricing decision. Sherlock's questionnaire data confirms this: when asked if he uses any price per gram method Sherlock answered in the negative and went on to remark that for him mass is completely delinked from his pricing decisions. This supported the removal of the mass attribute from his multicriteria model.

Sherlock was satisfied with the strategic map (Figure 6.6), but wanted to make some modifications to the location of the breakpoints. For example, with the age attribute Sherlock believed firmly that there are only two time-periods of interest: the period before 1820 (i.e. from 1750 - 1820) and post (i.e. from 1820 - to cessation around 1885). Consequently his interaction consisted of a mixture of piecewise linear approximation and bisection methods.

Figures 7.2, 7.3 and 7.4 summarise the questions relating to makers. For Figure 7.2 focusing on maker's renown, we expect the number of makers in breakpoints one and two to be roughly equal (both are 15% of the whole distribution – See Table 7.1). Indeed, both of these are roughly the same size. The breakpoints three and four should contain slightly less makers than breakpoints one and two, with 10% of the distribution. Breakpoint four proportionally contains the correct number with respect to breakpoint one and two, but breakpoint three is relatively over populated. In this instance breakpoint zero should contain 50% of the distribution, but only contains one maker. Two reasons arise to explain this: firstly because the makers chosen were informed by the data which Sherlock had priced, and hence that data was systematically biased away from the makers which appear very infrequently; and secondly because some makers are not yet known.

The aforementioned lack of makers in breakpoint zero (of maker's renown), can also be partially understood by examining Figure 7.4 which reflects the maker's rarity of output. Table 7.1 gives each level as consisting equally of 20% of the distribution. Figure 7.4 shows that the number of makers in these breakpoints are roughly equal except for the last breakpoint. It is the makers in the last level (level four) whose output is very scarce (thus systematically not included), which

[2]For this type of collector the interest would be in the different categories: building up a collection of different objects from the different time periods representing makers and styles.

[3]This collector would be interested in objects produced by ancestors only.

[4]All Table Silver (e.g. forks, spoons, ladles) is of size classification small.

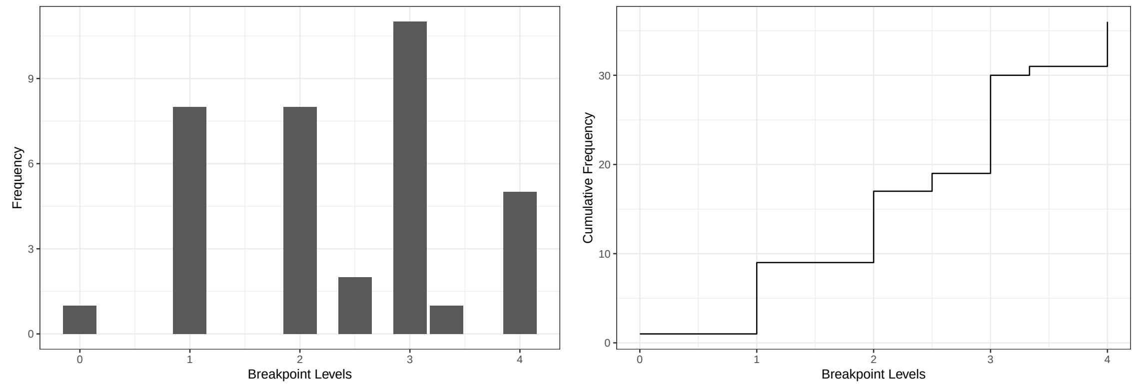


Figure 7.2: Sherlock's distribution for maker's renown. The right-hand figure is a cumulative density.

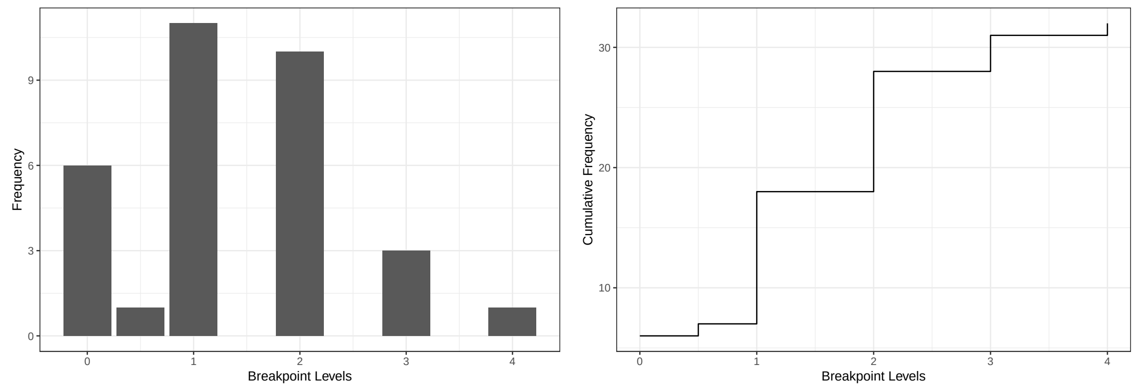


Figure 7.3: Sherlock's distribution for maker's craftmanship. The right-hand figure is a cumulative density.

partially explains why there are not more makers in the zero level of maker's renown (Figure 7.2).

Finally, notice that Figure 7.3 has the maker's craftmanship distribution reasonably close to that theorised in Table 7.1 for each of the levels. Once more there are relatively few makers in level zero, but the reasoning for this artefact has already been explained.

Now to focus on Sherlock's multicriteria model. Figure 7.5 presents Sherlock's partial value functions. Table 7.6 gives the score response to the different categories of Cape silver. Table 7.7 gives the weights for each of the partial value functions.

The sharp increase in score between the last two levels' breakpoints for maker's renown (Figure 7.5) reflects that demanding a larger collector interest is easier when constrained by a text description (with accompanying pictures) for the most well known makers (rather than less well-known ones). With the exception of this attribute and the maker's craftmanship attribute, Sherlock's partial value functions are close to linear. Sherlock's maker's craftmanship partial value function includes some points determined using bisection methods, since not all levels (as Figure 7.3 demonstrates) were whole numbers. This arose as there were check cards (for the same maker) where the DM assigned different levels and the average was taken.

Table 7.6 gives the detail for the different scores of each category. The categories (that Sherlock rated) were chosen to reflect the items Sherlock sells. Notice that it is overwhelmingly Table Silver – each subgroup of Table Silver in fact has a different score. It is clear that this moves away from the theoretical divisions of rarity that previously occurred in Chapter 6. The current direct evaluation of the different subgroups was easier.

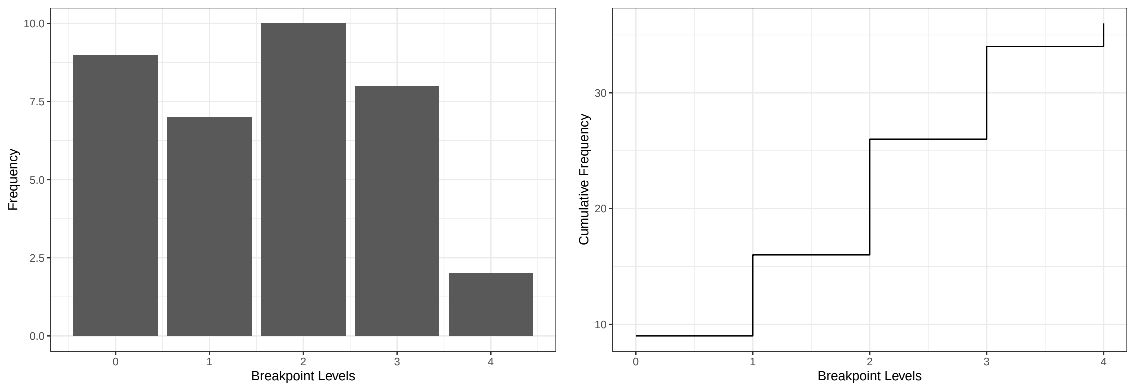


Figure 7.4: Sherlock's distribution for maker's rarity of output. The right-hand figure is a cumulative density.

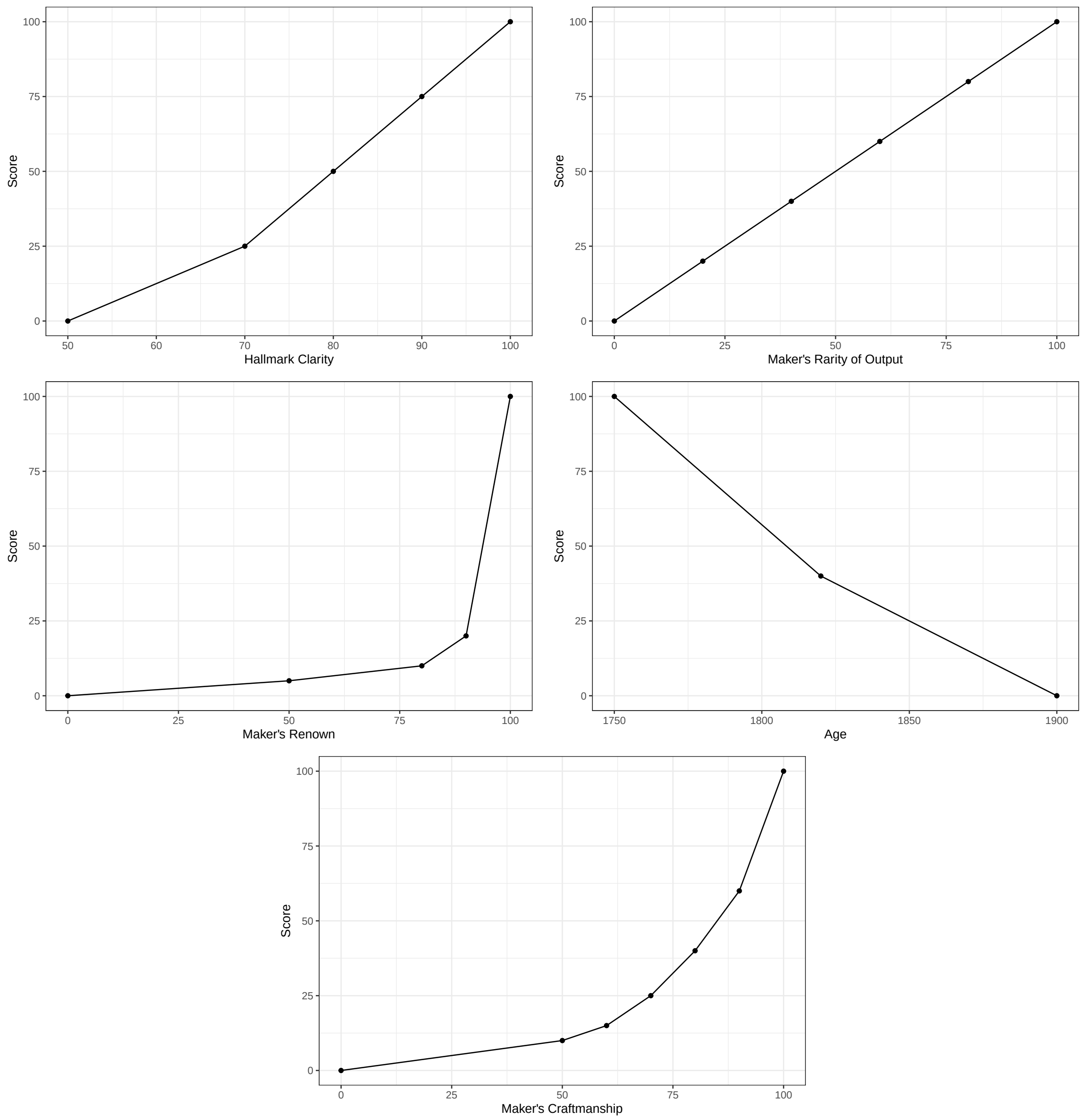


Figure 7.5: Sherlock's partial value functions.

Table 7.6: Sherlock's scores for each of the categories of Cape silver

Category	**Score**
Table Silver- All other	0.0
Table Silver- Spoons- Salt Spoons	30.0
Table Silver- Spoons- Mustard Spoon	45.0
Table Silver- Knives- Butterknife	50.0
Table Silver- Ladles	55.0
Table Silver- Forks- Konfyt	57.5
Table Silver- Marrow Scoops	60.0
Table Silver- Spoons- Orange Spoons	60.0
Table Silver- Spoons- Basting	65.0
Drinking Vessels- Mugs- Christening	67.5
Table Silver- Fish Slices	72.5
Table Silver- Sugar Nips & Sugar Tongs	83.5
Table Silver- Napkin and Place Holders	85.0
Boxes- Snuff Boxes	90.0
Table Silver- Knives- Tableknife	95.0
Boxes- Vinaigrettes	100.0

Table 7.7: Sherlock's weights for each of the partial value functions. Weights were determined using MACBETH.

Partial Value Function	**Weight**
hallmark clarity	33.33
category of silver	27.28
maker's rarity of output	18.18
maker's renown	12.12
age	6.06
maker's craftmanship	3.03

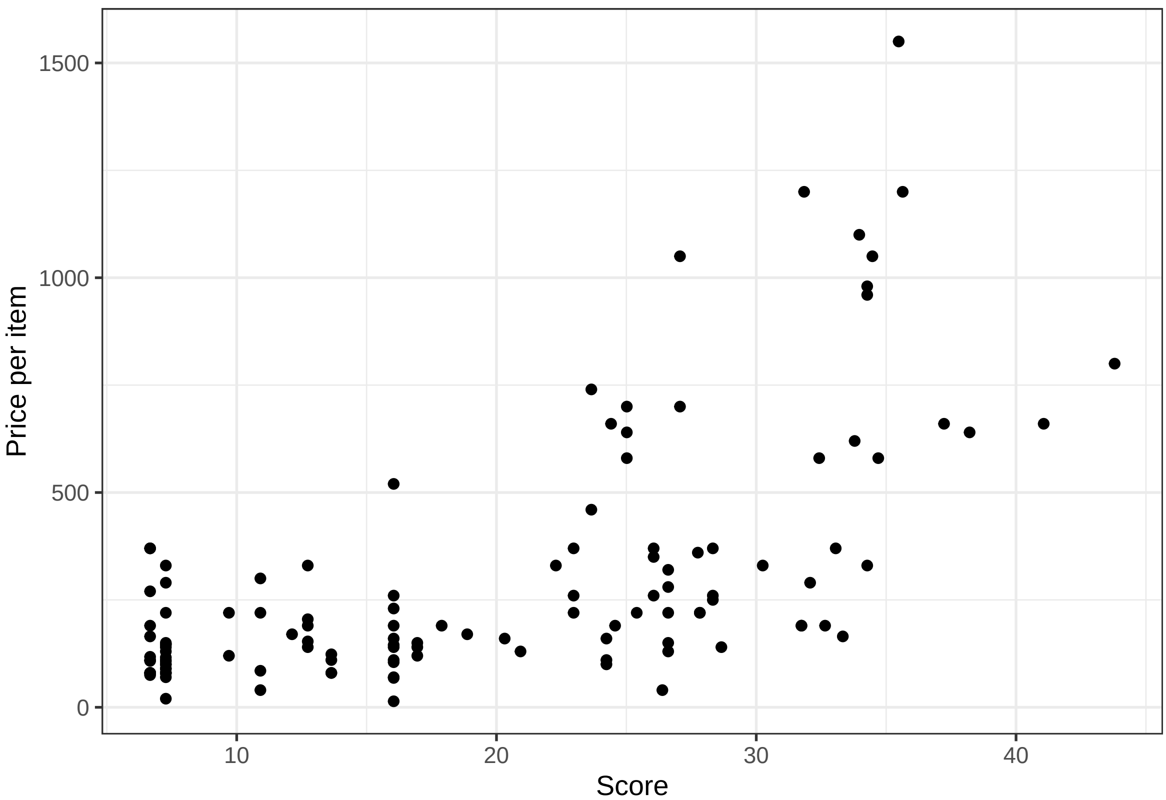


Figure 7.6: Sherlock's MCDA model's score and his assigned price per item. (n = 135)

Figure 7.6 shows Sherlock's multicriteria decision model's score evaluation for each item compared to a price per item that was previously assigned. No currency is nominated to protect Sherlock's anonymity. The data, $n = 135$, comes from his pricing over the period from 2009 until the present. There are less items with higher scores i.e. the distribution of scores is positively skewed, rather than symmetrical (see Figure 7.7).

Returning to Figure 7.6, the score evaluation (x-axis) is a value measure (the output of his multicriteria additive model). By construction this measure is well bounded (viz. in this specific case between [0,100]). Sherlock expected that as score increased so price would grow at a rate proportional to its current value, therefore suggesting non-linear returns[5].

Generally with utility functions the focus is on how the utility score $u(x)$ depends on the variable x. However, here the interest is reversed: what can $u(x)$, the score, tell us about price x. Two questions that need to (and will) be tackled: to what extent is price an expression of the utility score, and how to characterise this utility function. The next section is a brief interruption to the analysis of Sherlock to properly discuss utility and inverse utility functions.

[5] Sherlock explained that "exceptional pieces demand exceptional money" – by this he meant that pieces on the tail-end of the score distribution demand considerably more than those in the middle.

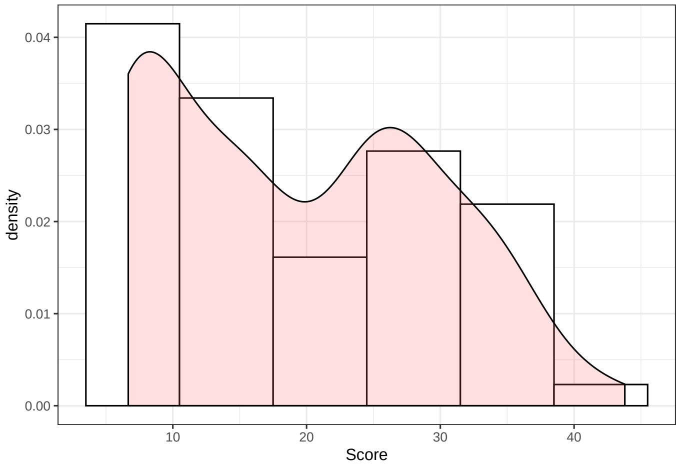


Figure 7.7: The distribution of Sherlock's multicriteria model score.

7.3 A Discussion on Utility Functions

The relationship between MCDM score (i.e. the value measure) and price is clearly increasing: as score increases so does the price. The utility function could be of the class risk-averse (Eq.(7.1)), risk-neutral (Eq.(7.2)) or risk-prone (Eq.(7.3)) (Keeney, 1996, 143):

$$u\left(x\right) = a + b\left(-e^{-cx}\right) \tag{7.1}$$

$$u\left(x\right) = a + b\left(cx\right) \tag{7.2}$$

$$u\left(x\right) = a + b\left(e^{cx}\right) \tag{7.3}$$

where $a, b > 0$ ensures that u is scaled from zero to one (or on the desired scale). For increasing utility functions we expect $c > 0$. Further, for the risk-neutral utility function (Eq.(7.2)) $c = 1$. In Eq.(7.1) and Eq.(7.3) the degree of the DM's risk aversion is given by c, suggesting that risk aversion can be compared.

These equations can be rewritten to highlight the value of x for different utility values u. Respectively the inverse utility functions are Eq.(7.4) (risk-prone), Eq.(7.5) (risk-neutral) and Eq.(7.6) (risk-averse):

$$x\left(u\right) = \frac{-1}{c}\ln(\frac{a-u}{b}) = \frac{1}{c}[\ln(b) - \ln(a-u)] \tag{7.4}$$

$$x\left(u\right) = \frac{u-a}{bc} \tag{7.5}$$

$$x\left(u\right) = \frac{1}{c}\ln(\frac{u-a}{b}) = \frac{1}{c}[\ln(u-a) - \ln(b)] \tag{7.6}$$

and the following restrictions apply: Eq.(7.4) $u < a$ and Eq.(7.6) $u > a$.

Consider Figure 7.8 which has its u bound between [0,1] and x between $[x_0, x^*]$. Figure 7.9 provides the inverse utility which correspond to Figure 7.8. Note that the corresponding inverse functions are in reverse order to the original utility functions.

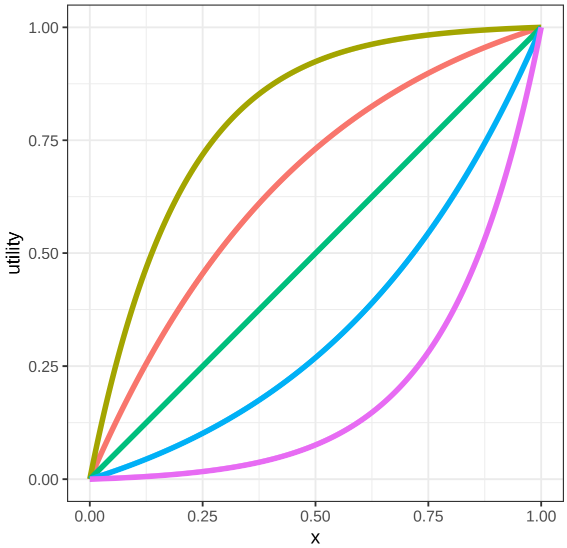


Figure 7.8: Various utility functions (from left to right): more risk averse (lime), risk averse (orange), risk neutral (green), risk prone (blue), more risk prone (purple).

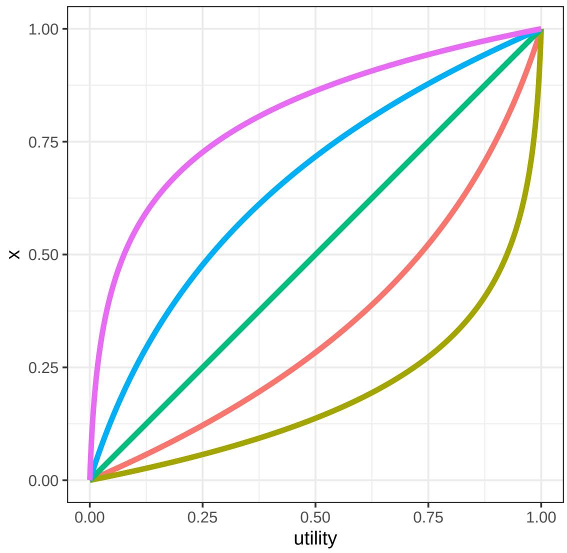


Figure 7.9: Various inverse utility functions (from left to right): more risk prone (purple), risk prone (blue), risk neutral (green), risk averse (orange), more risk averse (lime). Note the order change.

Comparing Figure 7.6 (which gives Sherlock's inverse utility function) and Figure 7.9, it is clear that Sherlock has a risk averse reward between price per item and utility.

To determine more about the relationship it is necessary to estimate a best fit: in this case one that minimises the mean squared error of $\hat{p} - p$, where p is the actual price per item and $\hat{p}$ is the predicted price given by

$$\hat{p} = f(\text{score}) = \frac{1}{c}[\ln(b) - \ln(a - \text{score})], \tag{7.7}$$

which is defined from Eq.(7.4).

When fitting any non-linear model it is necessary to supply initial conditions. While the ultimate objective is to provide an estimate of the parameters in Eq.(7.7), it is easier to start with a non-linear estimation of Eq.(7.1) because a simple (but realistic) assumption simplifies the number of parameters needing guesses for initial conditions.

Assuming that $u = 0$ when $x = 0$ (which is true, at least approximately, in most cases) it is easy to see that:

$$0 = a + b(-e^{-c\times 0}) = a - b \tag{7.8}$$
$$a = b. \tag{7.9}$$

Hence for the largest observed value of u (denoted by $u_{\max}$),

$$u_{\max} = b + b(-e^{-cx}) \tag{7.10}$$
$$u_{\max} = b(1 - e^{-cx}) \tag{7.11}$$
$$\frac{u_{\max}}{(1 - e^{-cx})} = b = a. \tag{7.12}$$

It is straightforward to visualise different guesses for the value of c comparing the predicted values against the data. The best guess can then be provided as an initial condition to a non-linear estimation routine. It is very important to use an estimation routine which respects bounds such as Elzhov et al. (2016) because unconstrained estimations tend to violate the condition that $u < a$. Even when having a bound the solution tends towards the imposed lower bound. The following are therefore required for estimating an equation in the form of Eq.(7.1): i). a lower bound $a = u_{\max} +$ tolerance, where the tolerance ensures a potential solution is not too close to $u_{\max}$; ii). a lower bound on b, c such that they are greater than zero (in order to respect the theoretical restrictions). The final parameter estimates from the above procedure make excellent initial conditions for estimation of Eq.(7.4), which is also done using a non-linear estimation routine and should include the same bounds as before.

7.4 RESPID 1: Sherlock (continued)

Figure 7.10 (page 75) builds on Figure 7.6 (page 70), adding the best fitting inverse utility function. Eqs.(7.13) and (7.14) give Sherlock's utility function and inverse utility function respectively.

$$u(x) = 50.97 + 54.14\left(-e^{-0.00201x}\right) \tag{7.13}$$
$$x(u) = \frac{1}{0.00201}[\ln(54.14) - \ln(50.97 - u)] \tag{7.14}$$

Sherlock is a dealer/retailer which should be remembered when trying to understand why he has a risk-averse utility function. Perhaps high ticket items may be slow moving. Expensive items also absorb capital that may have otherwise gone towards multiple different objects. Sherlock likely prefers more items with lower returns, but manageable risks; rather than stock that may have high returns, but could also involve protracted time until sale. Although not seen as clearly with Sherlock, it is worth mentioning that the next DM – Holmes – considers the economy as a highly important tribute. Shifts in the economy could change an item's value dramatically – but more on this and links with risk aversion will be discussed in Sections 7.5 and 8.1.

The utility (and inverse utility) function for a collector may look different to that of a dealer/retailer. Perhaps more risk taking would be evident: reflecting that collecting is not their livelihood, but a hobby. In juxtaposition with the retailer, from the collector's perspective it could be that having one good (i.e. high scoring) piece is better than having several lessor (i.e. lower scoring) pieces, even if the total expenditure is the same. These threads are explored further in the future research section of the conclusion.

On closer examination of Figure 7.10, there is evidence of unaccounted 'spread': there are objects with the same score but a large price range (around score = 16 for example). In reviewing the data, one of the causes for this is that it is essentially the same piece (for example another spoon by Lawrence Twentyman), but the currency value has progressively increased over time. Unfortunately, since the data spans 2009 until the present, it is difficult to de-couple the effect of inflation and the appreciation of the objects. Should all the data be repriced in a single sitting it may be likely that the revised points which currently form a vertical line would be clustered closer together.

Secondly, (as alluded to earlier in this subsection) Sherlock's model consist only of attributes relating to the silver objects themselves. Holmes's model goes further – directly incorporating variables like the economy into the score. Shifts in the economy could account for changes in price without a current change in score.

A third reason could be due to the existing problem with pricing such goods: the DM has inherent variability in what price he would give the object. Sherlock has already described himself as not being able to determine exactly the same price for an object two days apart. The variations in price around the same score as seen in Figure 7.10, could be realisations of this.

Addressing these three sources of spread provide an avenue for further research. One method that could tackle both the first and third problems would be to use a combination of both real and simulated data. The conclusion has suggestions regarding how to tackle this using simulated data in specially designed arrays, Chapter 9. Dealing with the second source of spread by incorporating market level variables will be looked at shortly.

Sherlock's model is rich in insight, however there are limitations. A key limitation that should be noted is that Sherlock feels his elicited values would differ, perhaps dramatically, under repeated elicitations. This suggest that a single session may not be sufficient to create a complete MCDM. Sherlock's model probes slightly the idea of a value or knowledge based decision support system to guide the DM on prices of objects for which there is no historical data. However, Section 8.1 implements a price based decision support system and explains how, in addition to price estimates, the DM can use such a system to learn more and refine their value judgements. A second type of decision support system could be a collection of Sherlock's MC models built up over years. The current MC model would form the first contribution to this. Having a collection of MC models would allow one to track – through re-weighting and adding additional criteria – how judgements change over time. A historical analyst interested in silver would use such a decision support system to gain key insights into the thoughts of antique dealers/retailers in a specific time period. The Antique Dealers Research Project, based at the University of Leeds, United Kingdom, studies "the history of the antique trade during the last 200 years. The project is international in scope [. . . and seeks..] to explore key questions related to the history of 'antiques'"(Antique Dealers Research Project, 2020). To such a researcher MC models provide insights that were not possible to achieve using only ethnographic techniques.

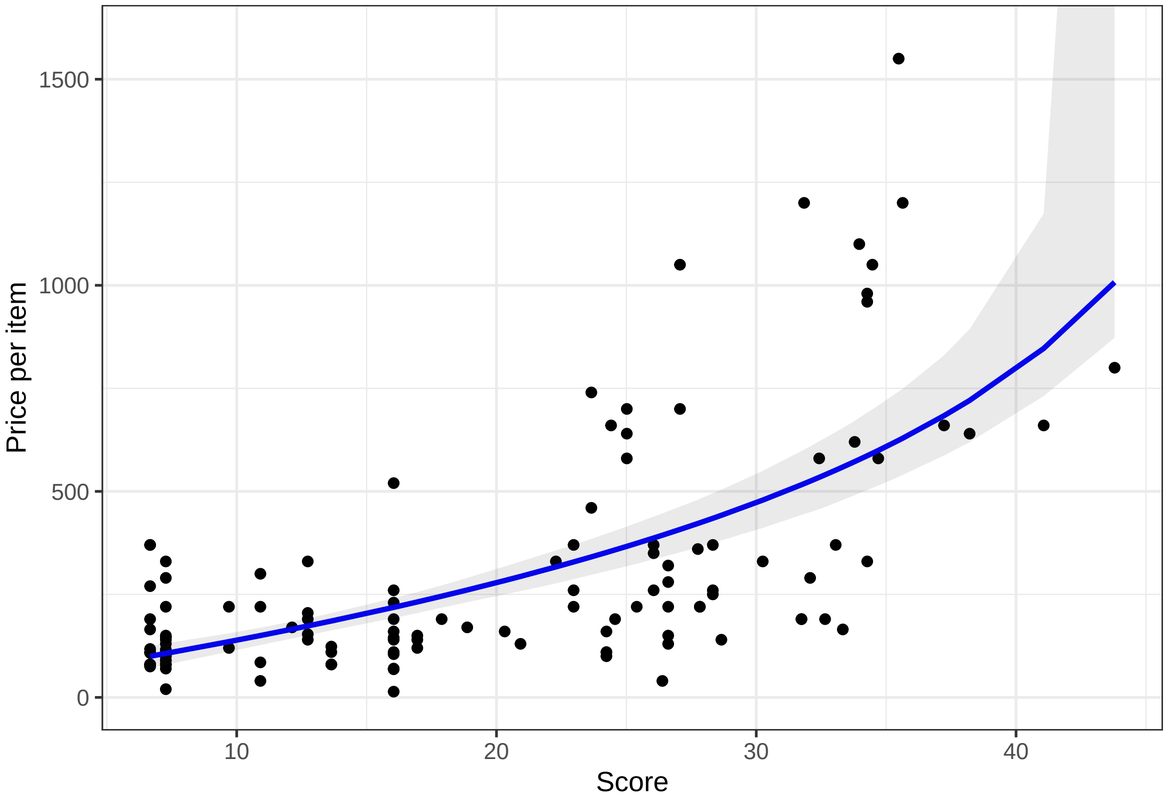


Figure 7.10: Sherlock's MCDA model's score and its relationship to price per item. A risk-averse inverse utility function provided in blue with a 95 percent confidence interval given by the grey region.

7.5 RESPID 2: Holmes

Holmes operates a successful business selling Cape silver objects of all sizes and prices. He regularly consults for other businesses and actively participates in the Cape silver community. Holmes' initial interview lasted a total of six hours. Thereafter various refinements occurred. His multicriteria model can be considered the richest yet discussed.

Holmes takes a very comprehensive approach to considering the attributes which affect collectors' interest and consequently have some relationship with price. During the problem structuring stage of the interview, Holmes expanded on the revised problem structuring map as depicted in Figure 6.6. His problem structuring map (Figure 7.11) includes four new attributes: the appearance of new reference material (new information), the economy, entry of a new collector and whether the object has a particular presentation inscription.

The appearance of a new reference (for example a new book with original research) or a new research paper, Holmes explains, spurs on collectors' interest because it creates a talking point. The economy, which Holmes currently considers to be neutral, also affects collectors' interest. When the economy is more buoyant there is greater amount of free cash (including from corporates) available to make purchases; conversely when the economy is down, there is less collectors' interest. The entry of a single new collector in such a specialised market also makes an impact: a new collector with a sizable interest can totally reshape a market, Holmes explains. Lastly, a presentation inscription differs from an engraving (which shows ownership) by having important historical ties. For example if the Cape Governor presented a silver object to someone and motivated the reasoning for this presentation on the object, this would count as a presentation inscription.

Now to focus on Holmes's multicriteria additive model. Figure 7.12 gives all of these new, as well as the other familiar, partial value functions. Table 7.8 gives the category attribute and demonstrates the breadth of Holmes' business. The number of subcategories is large. The weights for the partial value functions are given by Table 7.9. For the age attribute the bisection method was used. Holmes included modern silver as part of his definition of Cape silver, hence the final breakpoint for this attribute is at the year 2020. However, when the weight of the age attribute is considered (Table 7.9) in conjunction with the partial value function, it becomes apparent that age's contribution is really negligible to the overall score. Holmes' partial value functions are mostly linear. Unlike Sherlock who places large weight on hallmark clarity, Holmes remarks: "They come, look, stare and are taken by the object as a whole". Such a remark corroborates the richness of the model in that it can sum up both of the DMs thoughts, and how their businesses are positioned.

Holmes had no readily available price data. In order to relate his multicriteria model's score to price it was necessary to create (simulate) a sample of data and then ask Holmes to provide the price. The images and descriptions used to simulate this data was drawn from historical records and Holmes' stock. This meant that a sample of both lessor and more remarkable pieces (many of which Holmes has in fact sold in the last 10 years) were considered. The process of simulating a dataset in this way has some advantages in that the data is considered at the current time period for the given market variables (economy and presence of new information). However it does not overcome the internal inconsistencies that may be present. Returning to Holmes the next day with the same dataset may indeed yield different prices. One way of tackling this would be to elicit repeated pricings of the same object from DMs over different days in a close time period. This is discussed more in the further research section of the conclusion.

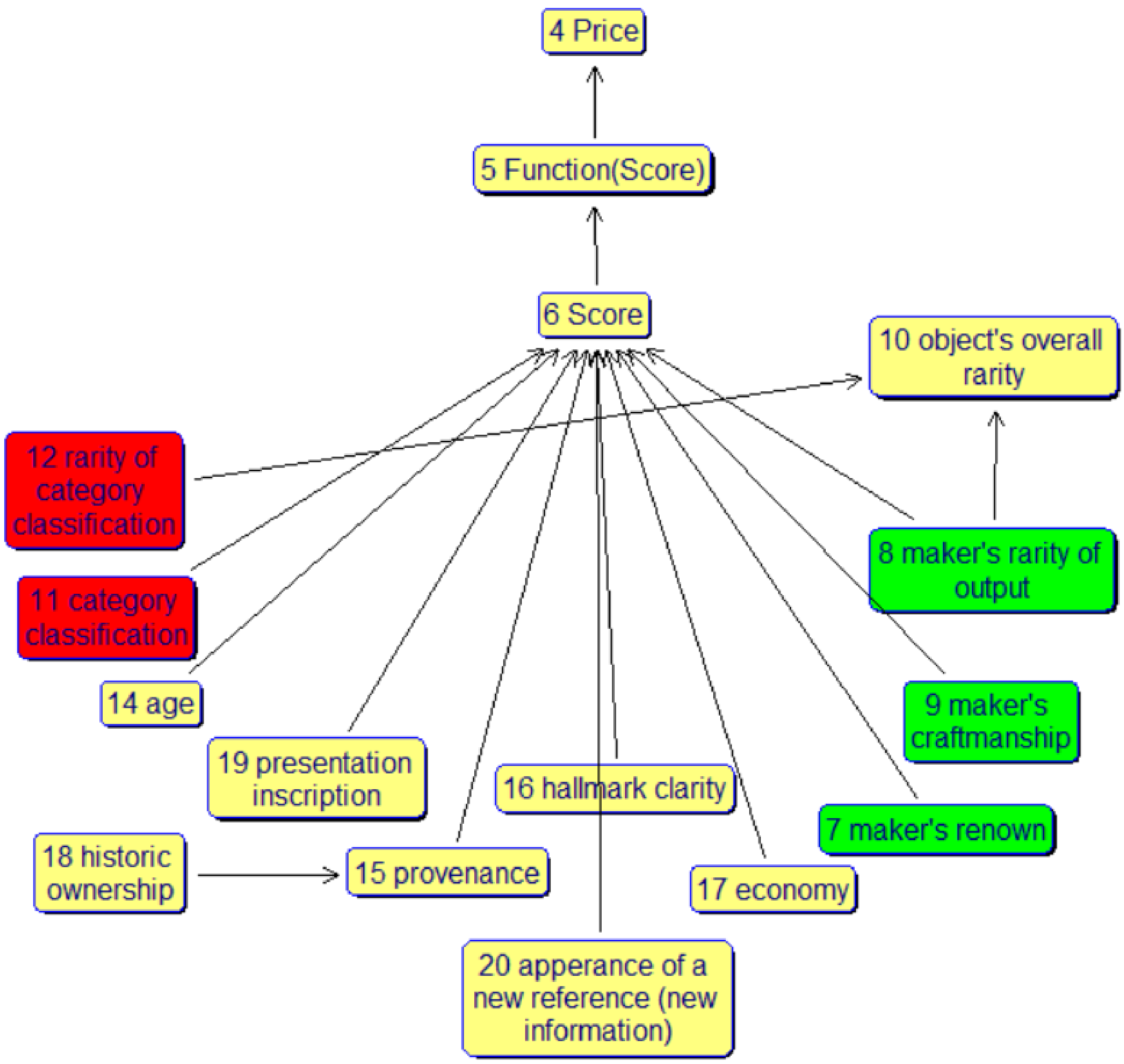


Figure 7.11: Holmes' Revised Strategic Map

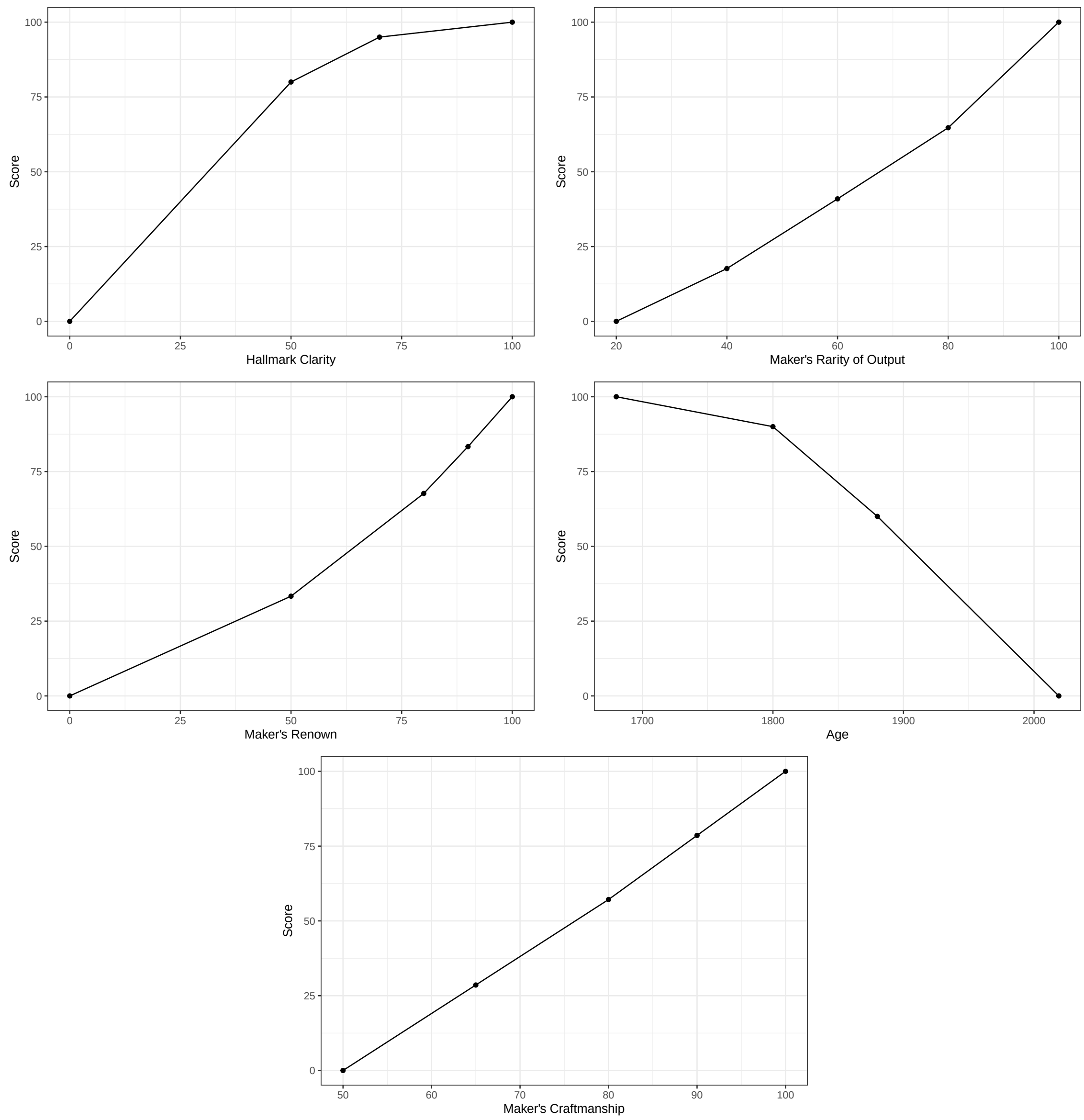


Figure 7.12: Holmes' partial value functions.

Table 7.8: Holmes' scores for each of the categories of Cape silver

Category	**Score**
Vase	250.0
Candlestick	200.0
Tessie	200.0
Sugar Box	180.0
Beaker- Double Beaker	175.0
Brandy & Other Saucepans	175.0
Coffee Pot	175.0
Cups and Covers	175.0
Salvers & Trays- Salver	175.0
Walking Stick	170.0
Beacker- Beacker	150.0
Presentation Staff	150.0
Sugar Bowl	150.0
Tea Pot	125.0
Wine- Labels	100.0
Drinking Vessels- Mugs- Christening	90.0
Boxes- Snuff Boxes-Cornie	72.5
Boxes- Vinaigrettes	70.0
Table Silver- Marrow Scoops	70.0
Boxes- Snuff Boxes	65.0
Table Silver- Ladles	60.0
Bible Clasp	50.0
Table Silver- Knives- Tableknife	45.0
Table Silver- Spoons- Orange Spoons	45.0
Table Silver- Fish Slices	40.0
Table Silver- Spoons- Basting	40.0
Table Silver- Forks- Konfyt	35.0
Table Silver- Spoons- Salt Spoons	35.0
Table Silver- Sugar Nips & Sugar Tongs	35.0
Table Silver- Knives- Butterknife	30.0
Table Silver- Spoons- Mustard Spoon	25.0
Table Silver- Spoons- Tablespoons	5.0
Table Silver- All other	0.0
Table Silver- Napkin and Place Holders	

Table 7.9: Holmes' weights for each of partial value functions. Weights were determined using MACBETH.

Partial Value Function	**Weight**
presentation inscription	16.01
provenance	16.01
new collector	13.58
category of silver	13.24
maker's craftmanship	11.18
economy	10.78
maker's rarity of output	6.96
maker's renown	4.89
new information	4.54
hallmark clarity	2.45
age	0.36

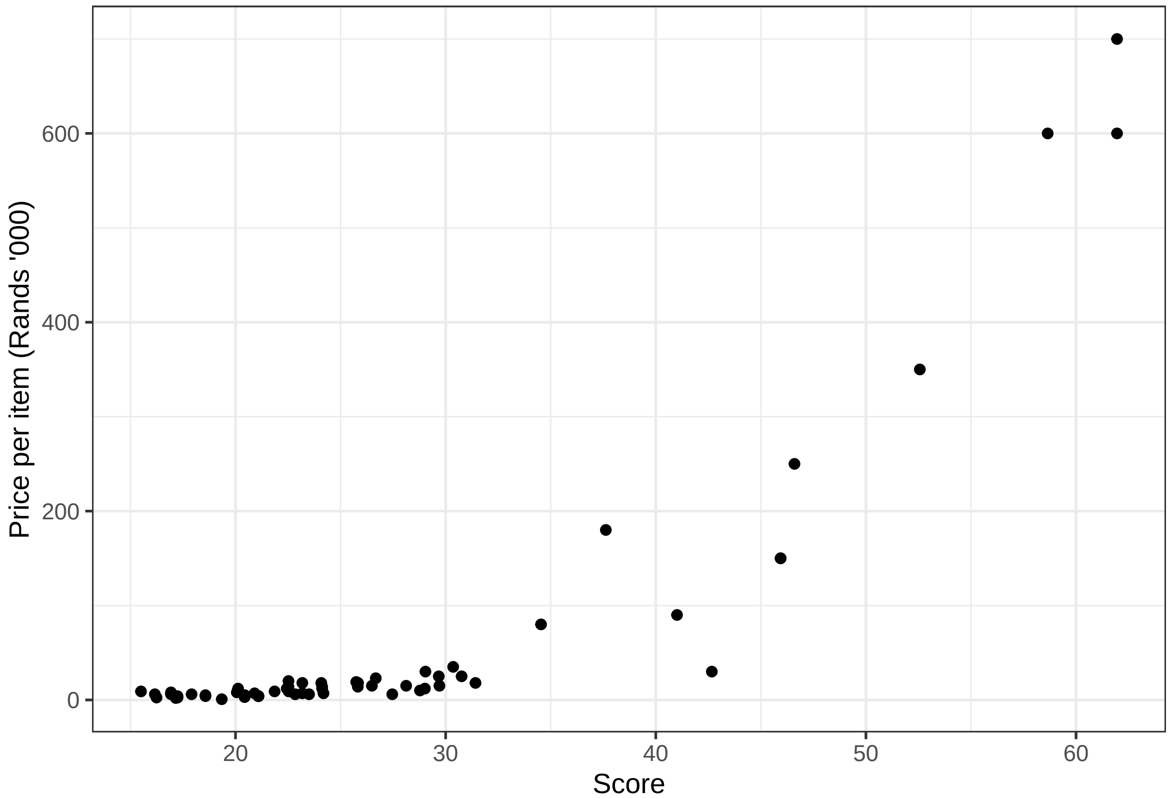


Figure 7.13: Holmes' MCDA model's score and its relationship to price per item.

Figure 7.13 (page 81) gives the relation between multicriteria model score and price for Holmes. One immediately remarkable feature of this graph is the range. Prices, divided by item, vary from as little as R750 all the way to R700'000. Between the price of R20'000 and R700'000 there are only a few observations. This reflects the fact that although simulated data for price was used, the object's characteristics where drawn from historical auction records. Since the 2000s only these few high-price pieces have occurred.

The density distribution in Figure 7.14 shows that score is positively skewed. Presentation and Provenance, although item specific, occur in very few of Holmes' cases. The first significantly weighted attribute which occurs for all items (and is item specific) is the category. Table 7.9 gives the attribute weights. Comparing Figure 7.14 and Figure 7.15: the expensive (high scoring) items are typically those whose category judgement exceeds 100 i.e. they exceed the aspiration level. (Table 7.8 also gives the category judgements.) As with score, the distribution of prices is also positively skewed.

Before Figure 7.13 was available, Holmes was asked to sketch how he thought his model's score would relate to price. (Bearing in mind that the calculations which allowed Holmes's assessment of score for the items still needed to be completed.) The sketch that Holmes drew is very similar in shape to that of Figure 7.13 and this, Holmes remarked, pleased him greatly.

Figure 7.13 suggests a risk-averse value function. Fitting a risk-averse value function was an interesting engagement. Figure 7.16 provides two such functions, one given in red, the other in blue. The red is the best fitting line generated from risk-averse utility function equations given in Section 7.3.

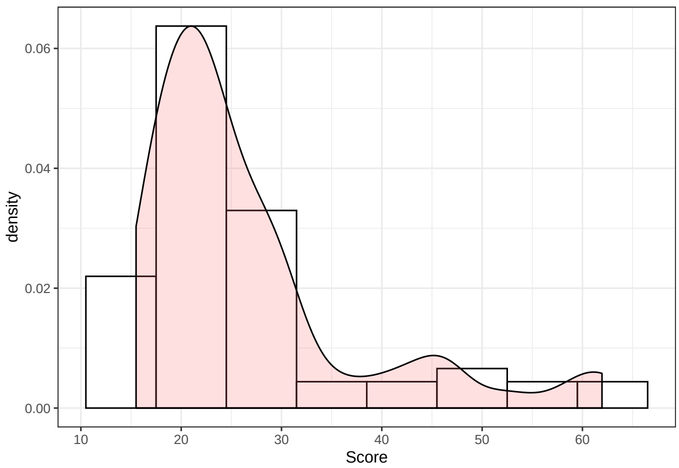


Figure 7.14: The distribution of Holmes's multicriteria model score.

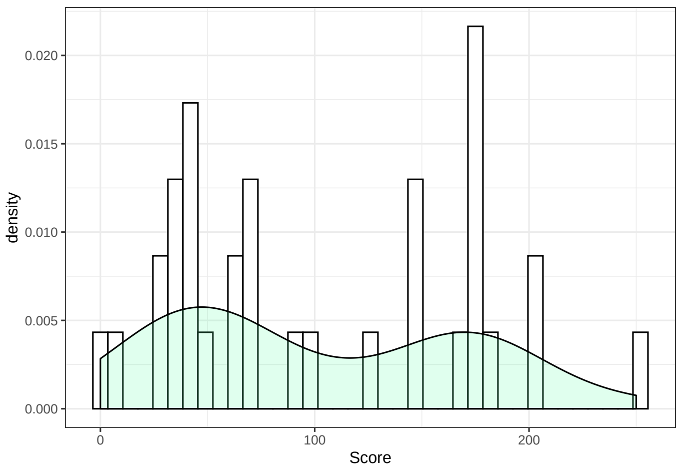


Figure 7.15: The distribution of Holmes's category judgement values.

Eqs.(7.15) and (7.16) give Holmes' utility function and inverse utility function (the red line) respectively.

$$u\,(x) = 64.24 + 41.11\left(-e^{-0.00389x}\right) \tag{7.15}$$

$$x\,(u) = \frac{1}{0.00389}[\ln(41.11) - \ln(64.24 - u)] \tag{7.16}$$

From inspecting Eq.(7.16) graphically in red on Figure 7.16, it is clear why Holmes was dissatisfied with it. For many objects with small values of score (roughly below score=25), the resultant price would be negative. The reason why this function fits as it does comes from the constraint that the parameter $a > u$. Relaxation of this criteria results in a fit that is better visually (as well as has a smaller residual sum of squares); but then does not generate a valid risk-averse (exponential) utility function. This is an interesting dilemma. It points to the fact that other classes of utility functions should be explored – an avenue for further research.

How best then to capture Holmes' relationship between score and price? The analyst and the DM undertook a several different iterations until Holmes settled on the relationship generated by the blue line. This function (the blue line) basically maintains the shape of a risk-averse inverse utility function, while the price remains positive for small values of score. Although the blue line, generated using a smoothing function, (in this case a local polynomial regression fitting), does not easy translate algebraically, Holmes felt it was a much better representation of his own idea for his inverse value function. The function has an essentially horizontal region (roughly between score 5-25) – crucial to Holmes's reasoning – around which score may increase and decrease without too much change in price. Another approach, worthy of further investigation, would have been to divide the MCDM model into two. One model for low-score, low-price objects and another for high-score, high-price objects. The logic for this becomes apparent when working with the knowledge based decision support system constructed from Holmes' value judgements. In particular, varying large-weighted parameters (such as economy) can have little effect on items with a relatively low score (the item would move along the horizontal section of the function mentioned above); while having a remarkable effect on objects which have a higher score. To demonstrate, Section 8.1.2 includes this scenario and further discussion can be found there.

Holmes described the different parts of his inverse value function. This is actually easier to see if looking at the value function drawn in the standard fashion, as in Figure 7.17. Although not illustrated in this Figure, the value function Holmes described will start to plateau at high-value prices. This would reflect the fact that even as score increases there is a ceiling on price (generally speaking, set at the record price). For score values with a range between 15 and 20 rather than wiggling around, Holmes envisioned this as a straight vertical rise.

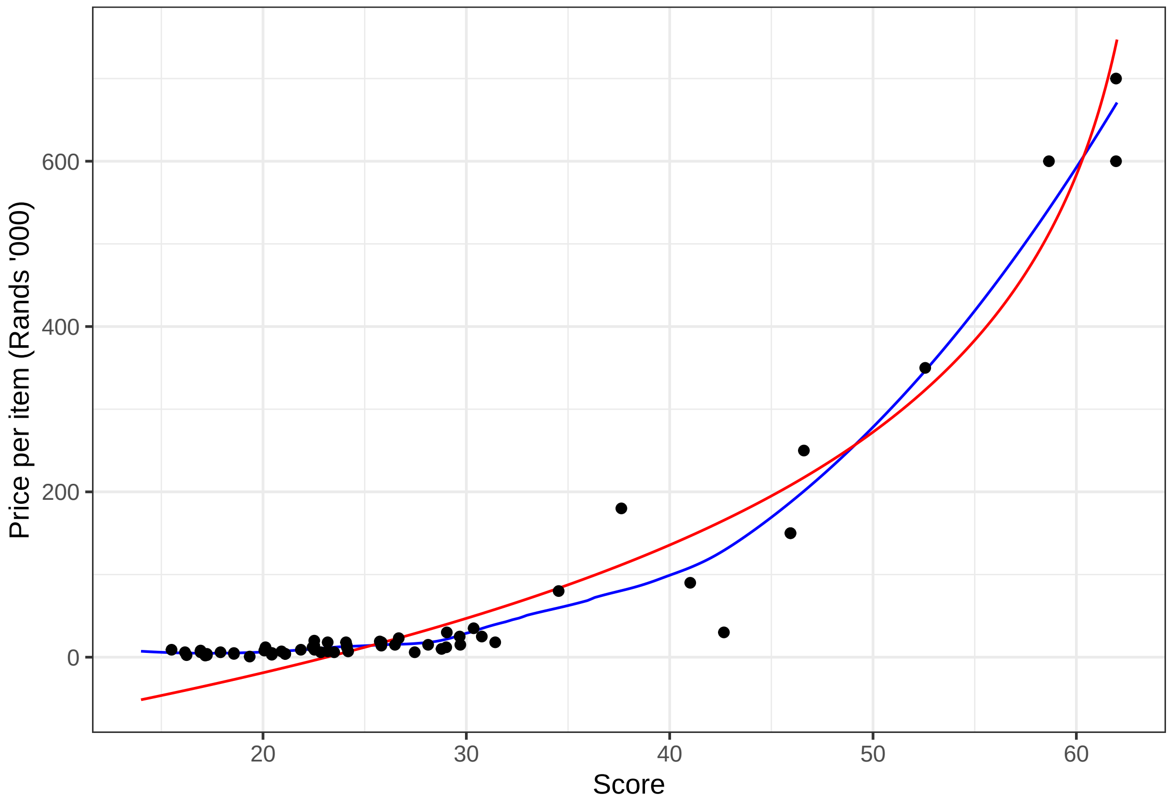


Figure 7.16: Holmes's multicriteria model scores with two potential candidate functions as inverse utility functions.

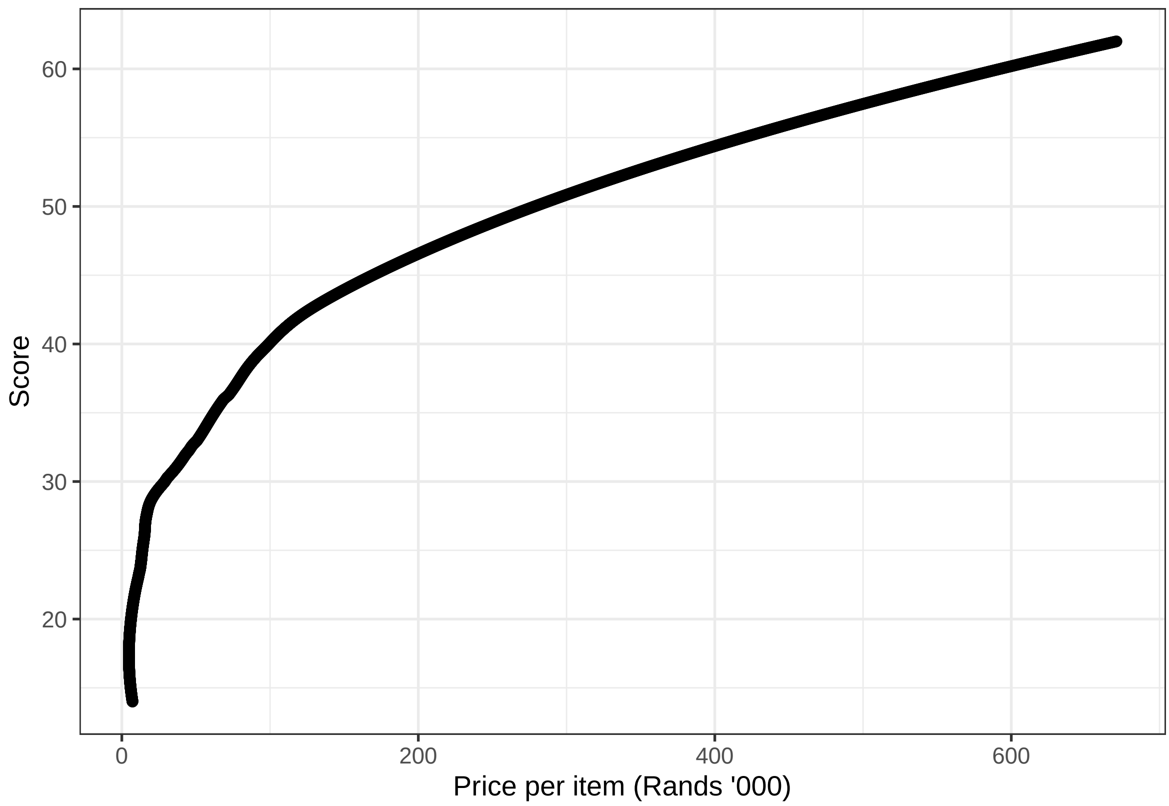


Figure 7.17: Holmes's utility function for price per item.

Further refinements of Holmes model may be possible. Holmes described an additional feature that his multicriteria model had not captured: overall object condition. The condition of the object does not change its underlying features (viz. its maker, its category, etc., nor even necessarily its hallmark clarity), yet it does effect the price. The condition of an object is judged relevant to the amount of degeneration that is expected over time. For example each time that silver is shined, in particularly improperly, some of the fine detail of the object is removed. Excessive and repeated shining can lead to a total loss of detail. Accidental damage (such as dropping an object) can lead to denting and to a silver collector this points to improper care of the object.

Silver objects being metallic are relatively easy to restore. Condition also encompasses consideration of restoration. In general restoration will improve the physical condition of the object, but the object is seen to be marred by the need for restoration. Large restorations such as the construction of a replacement part (eg. a finial for a teapot) require proper disclosure. There are many more aspects which can be discussed, but would not be appropriate for this dissertation.

To return to what is relevant Holmes explained is the question of how best to incorporate condition into the final model. Holmes indicated that the condition had a multiplier effect on the final score between [0.9, 1.1]. Figure 7.18 gives the location for the condition multiplier based off a modified Figure 7.11 (page 77), according to Holmes. Further research is needed to determine if this is indeed a sensible approach rather than including it directly as an attribute to the multicriteria model. Holmes explained though that a shortcoming with its direct inclusion is that in many cases this attribute would be towards the global maximum[6]. Investigating this is a point of further research discussed more in the conclusion, Chapter 9.

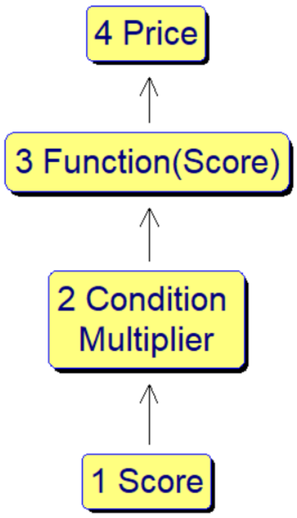


Figure 7.18: Modification to the problem structuring map showing the location of the condition modifier

[6]Many items perform well in the condition criteria.

Chapter 8

Decision Support Systems

A decision support system (DSS) is at its core a tool to support enhanced (informed) decision making. Generally DSS are implemented as computer software to ease calculation and allow for an improved user interface experience. DSS can take many forms, such as data-driven (uses historical data), knowledge-based (uses expert based rules and procedures) or model-driven (uses statistical models) (Power, 2002).

An important part of the multicriteria modelling process is realising an output that allows the DM to engage with their model. The multicriteria modelling process should allow a DM the structure in which to make enhanced decisions. In the case of repeated decisions such as a pricing decision (the design problematique), the modelling process does not yield a direct answer (as it would in the choice or sorting problematique), but rather a framework which needs to be engaged further.

This dissertation has resulted in two DSS, `CapeSilversmiths` (a data-driven DSS) described in Section 6.6, and another described here as the output of the modelling process. This DSS, `CapeSilverDecisions`, is a fully implemented web-deployed software with a carefully designed user interface intended to allow a DM who constructed a multicriteria model to engage and further refine their results.

8.1 Knowledge-Based Decision Support Systems

For illustrative purposes the multicriteria model for Holmes (RESPID 2) is implemented as a DSS named `CapeSilverDecisions`. Access to the DSS is an important consideration when creating one. In this instance the DM was consulted to find out how best to structure the DSS to suit his needs. Since this DM travels regularly, web deployment of the DSS was logical. Hence, the DSS `CapeSilverDecisions`, is accessible via the link https://capesilverdecisions.herokuapp.com/. Additionally, having an online DSS means that other role-players have easy access to its support.

A DSS designed to use a multicriteria model to estimate score is fundamentally interesting – not just to the DM whose model underpins it. With the lack of expertise in Cape silver such a device immediately becomes useful for other role-players: for example an auctioneer (outside of South Africa) who needs to price an object of Cape silver. The auctioneer and the DM come from the same class of role-player (those interested in price), but collectors can also find this DSS helpful because it explains the institution behind a field expert DM's thought process. Collectors can compare the DSS for Holmes and Sherlock and would be informed by their different views on the same problem. Insurers could use the different DSS to form aggregate opinions on replacement value. They can also justify a replacement value by referring to the attributes that influence its value. Individual sellers and small-scale retailers, who previously had a dearth of information, can now inform themselves better about possible value and which attributes of the object to highlight when advertising their wares. Many of these role-players would also be interested in the other DSS, the historical-based `CapeSilversmiths` (a data-driven DSS).

8.1.1 Interacting with `CapeSilverDecisions`

Upon opening the link in a web browser, the DSS loads and presents the user immediately with an operable screen. See Figure 8.1. The DSS includes all the attributes that Holmes had in his modelling process. These are divided into Market Variables (top row), Maker (top row) and Item Variables (middle row). The supporting calculations occur dynamically in the calculation table (bottom left) of the DSS.

The calculation table presents the attribute value, which is the DM's value judgement (score) for the particular level of the attribute. For example, for Holmes' a hallmark clarity of 50 corresponds to a score of 80 for this attribute. See Figure 7.12 (page 78) which gives all of Holmes partial value functions including that for hallmark clarity. The attribute's weights and resulting score (attribute values times weight) is also given. These features help explain the workings of the additive model to a user. Cumulative score keeps track of the contribution of each attribute to the final total score.

The final total score (i.e. the resulting value from the multicriteria additive model) is given below the graph of the inverse value function. The point estimate of the final score is shown graphically on the inverse value function graph.

Changing of a value such as selecting "Table Silver- Ladles" as the item category causes the DSS to update. Compare Figures 8.1 and 8.2 to see how the calculation table changes as well as the price estimates, and the graphics.

The language in this DSS has been kept intentionally simpler than would otherwise be used if this DSS were intended for a statistical audience; since, its users would come from different backgrounds. By making the mechanics explicit, it helps to introduce multicriteria modelling (and the additive value function) to the relevant role-players, and indeed the general public.

Knowledge-Based Cape Silver Decision Support

Decision Support | Table

About

Welcome to the knowledge-based decision support system for Holmes. This system can help you gain insights about potential prices of Cape Silver items. It uses Holmes's value judgements to construct a multicriteria value function and relates these judgments back to price. Many price estimates are indeed realistic, but for legal reasons etc., please consider this as a supplementary tool.

How to Use

Adjust the different attributes in the 'Decision Support' tab and observe how the score changes. The graphic shows your score-price estimate with a red dot. Depending on where you are on the graph sometimes a small change in score can result in a big price increase, while other times not. This is because of the exponential shape of the inverse value function. The calculated score and price are under the heading Final Calculations below the graph.

Learn More?

Read the workings in full in Jed Stephens's dissertation.

Market Variables

How is the economy?: Neutral

Has there been any new information or research related to Cape Silver? (Explanation to the right): None

As described by Holmes: 'A low impact source of information would be blog post or news paper article. A medium impact source would be a discovery of a new attribution for a maker or additional facts. A high impact source would be a new book or research output.'

Maker

Who made the item? Twentyman Lawrence Holme 1793

Item Variables

What category of item is this? Table Silver- All other

Does the item have any presentation inscription? No

For example it was presented to an important person or marked an important occasion.

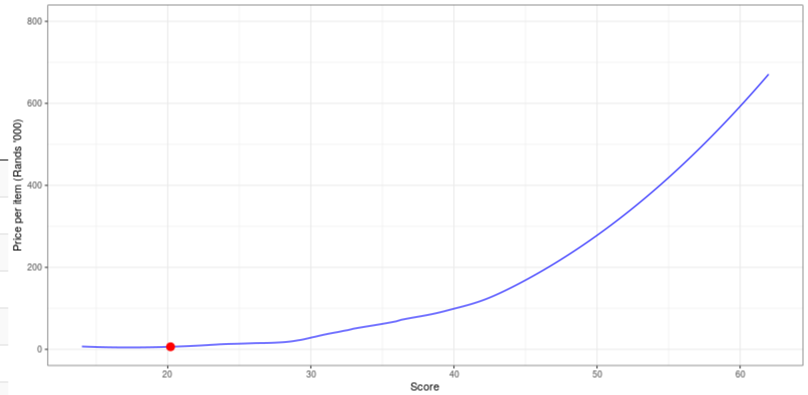
Does the item have provenance? None

Presence of a new collector (for this object type)? None

What is the clarity of the hallmark? (100 is perfect clarity) 0 – 50 – 100

How many items comprise? (Applicable only to Table Silver) 1 – 5

Attribute	Attribute Value	Weight	Score	Cumulative Score
presentation inscription	0.0	16.01%	0.00	0.00
provenance	0.0	16.01%	0.00	0.00
new collector	0.0	13.58%	0.00	0.00
category of silver	0.0	13.24%	0.00	0.00
maker's craftmanship	57.1	11.18%	6.38	6.38
economy	0.0	10.78%	0.00	6.38
maker's rarity of output	100.0	6.96%	6.96	13.34
maker's renown	100.0	4.89%	4.89	18.23
new information	0.0	4.54%	0.00	18.23
hallmark clarity	80.0	2.45%	1.96	20.19



Final Calculations

Final Total Score	Price (Rands)
20.19	6,284

Figure 8.1: Landing page for CapeSilverDecisions.

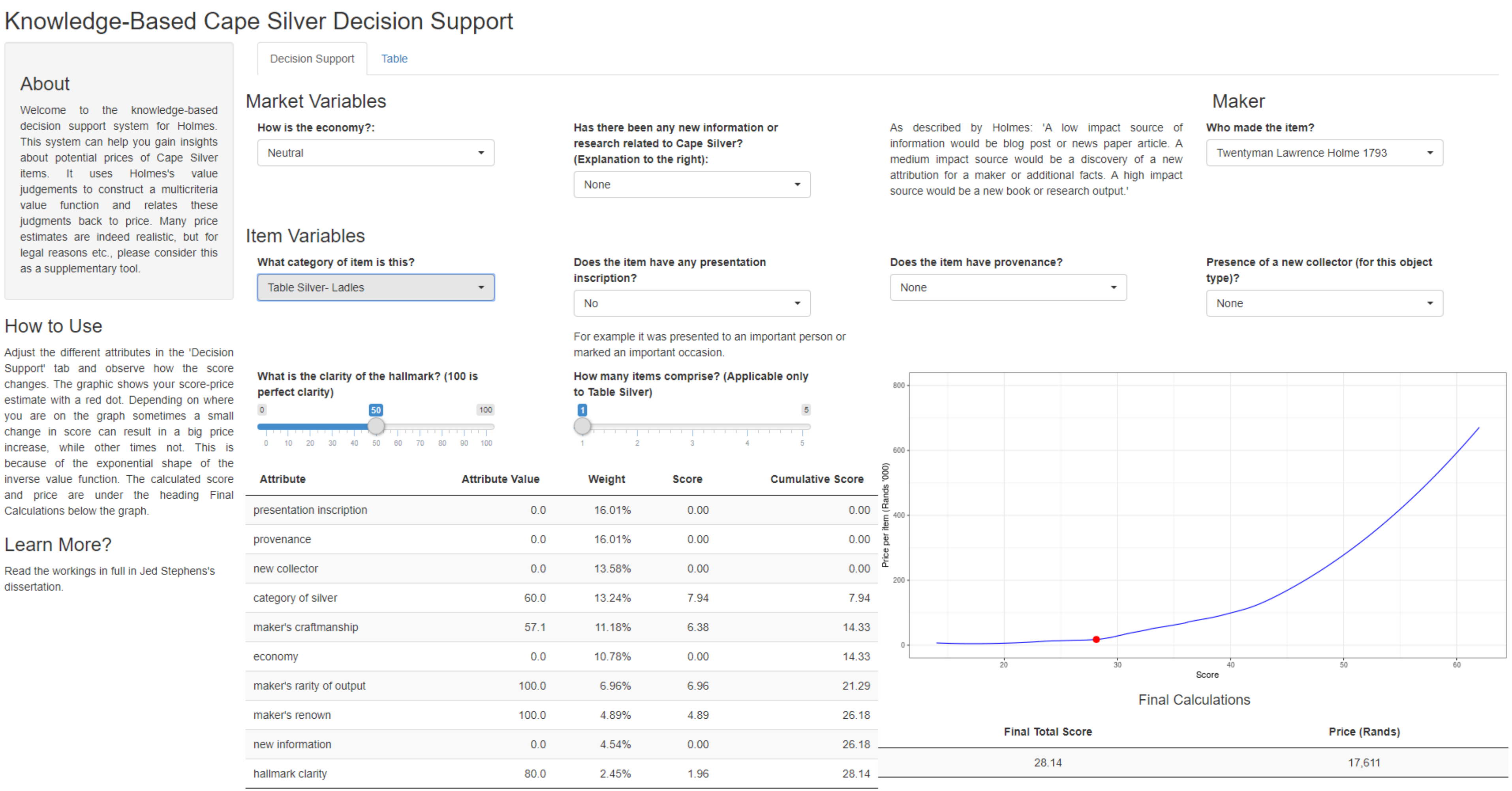


Figure 8.2: Pricing a ladle. A change of the category attribute from basic Table Silver.

8.1.2 Illustrating a DM's reasoning

Holmes described the use of the stable horizontal region in his inverse value function (extending from a score of 5 to 25) as ensuring that even high impact changes (for example the heavily weighted economy attribute) have low impact on certain categories of silver. (See Section 7.5 for details.) Figure 8.1 illustrates the lowest scoring category, basic Table Silver, made by a maker, Twentyman, who scores highly in all the maker attributes. The impact of an improved economy, moving to a level of "somewhat up" from "neutral" for the economy attribute, causes an additional 2.70 score points, changing the price by approximately R4000, Figure 8.3 illustrates.

In comparison, as Holmes intended, the price changes are more remarkable when moving from a higher base. For the category attribute, Figure 8.2 has the category ladle rather than the basic table silver in Figure 8.1, otherwise the objects are identical. The ladle object starts from a higher base score. The same score impact (caused by changing the economy), an additional 2.70 score points, now shifts the score out of the stable region into one of greater growth. The price for a ladle in that circumstance, as Figure 8.4 illustrates, increases by approximately R17'500. High scoring categories are best poised to take advantage of increases in score, but are also impacted more when the score declines. As we might expect, for less than "neutral" values in the economy attribute, the score declines. This behaviour illustrates the DM's experience of the market which is relatively stable for some categories of items, but not for others.

Knowledge-Based Cape Silver Decision Support

About

Welcome to the knowledge-based decision support system for Holmes. This system can help you gain insights about potential prices of Cape Silver items. It uses Holmes's value judgements to construct a multicriteria value function and relates these judgments back to price. Many price estimates are indeed realistic, but for legal reasons etc., please consider this as a supplementary tool.

How to Use

Adjust the different attributes in the 'Decision Support' tab and observe how the score changes. The graphic shows your score-price estimate with a red dot. Depending on where you are on the graph sometimes a small change in score can result in a big price increase, while other times not. This is because of the exponential shape of the inverse value function. The calculated score and price are under the heading Final Calculations below the graph.

Learn More?

Read the workings in full in Jed Stephens's dissertation.

Decision Support | Table

Market Variables

How is the economy?:

Somewhat Up

Has there been any new information or research related to Cape Silver? (Explanation to the right):

None

As described by Holmes: 'A low impact source of information would be blog post or news paper article. A medium impact source would be a discovery of a new attribution for a maker or additional facts. A high impact source would be a new book or research output.'

Maker

Who made the item?

Twentyman Lawrence Holme 1793

Item Variables

What category of item is this?

Table Silver- All other

Does the item have any presentation inscription?

No

For example it was presented to an important person or marked an important occasion.

Does the item have provenance?

None

Presence of a new collector (for this object type)?

None

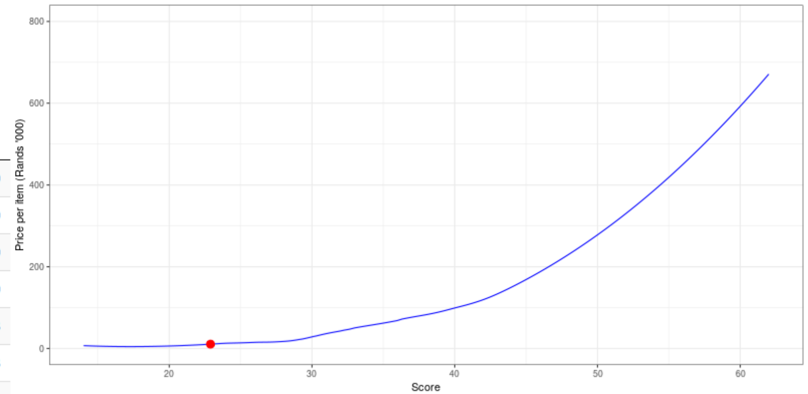
What is the clarity of the hallmark? (100 is perfect clarity)

0 — 50 — 100

How many items comprise? (Applicable only to Table Silver)

1 — 5

Attribute	Attribute Value	Weight	Score	Cumulative Score
presentation inscription	0.0	16.01%	0.00	0.00
provenance	0.0	16.01%	0.00	0.00
new collector	0.0	13.58%	0.00	0.00
category of silver	0.0	13.24%	0.00	0.00
maker's craftmanship	57.1	11.18%	6.38	6.38
economy	25.0	10.78%	2.69	9.08
maker's rarity of output	100.0	6.96%	6.96	16.04
maker's renown	100.0	4.89%	4.89	20.93
new information	0.0	4.54%	0.00	20.93
hallmark clarity	80.0	2.45%	1.96	22.89



Final Calculations

Final Total Score	Price (Rands)
22.89	10,645

Figure 8.3: Pricing table silver when the economy is somewhat up.

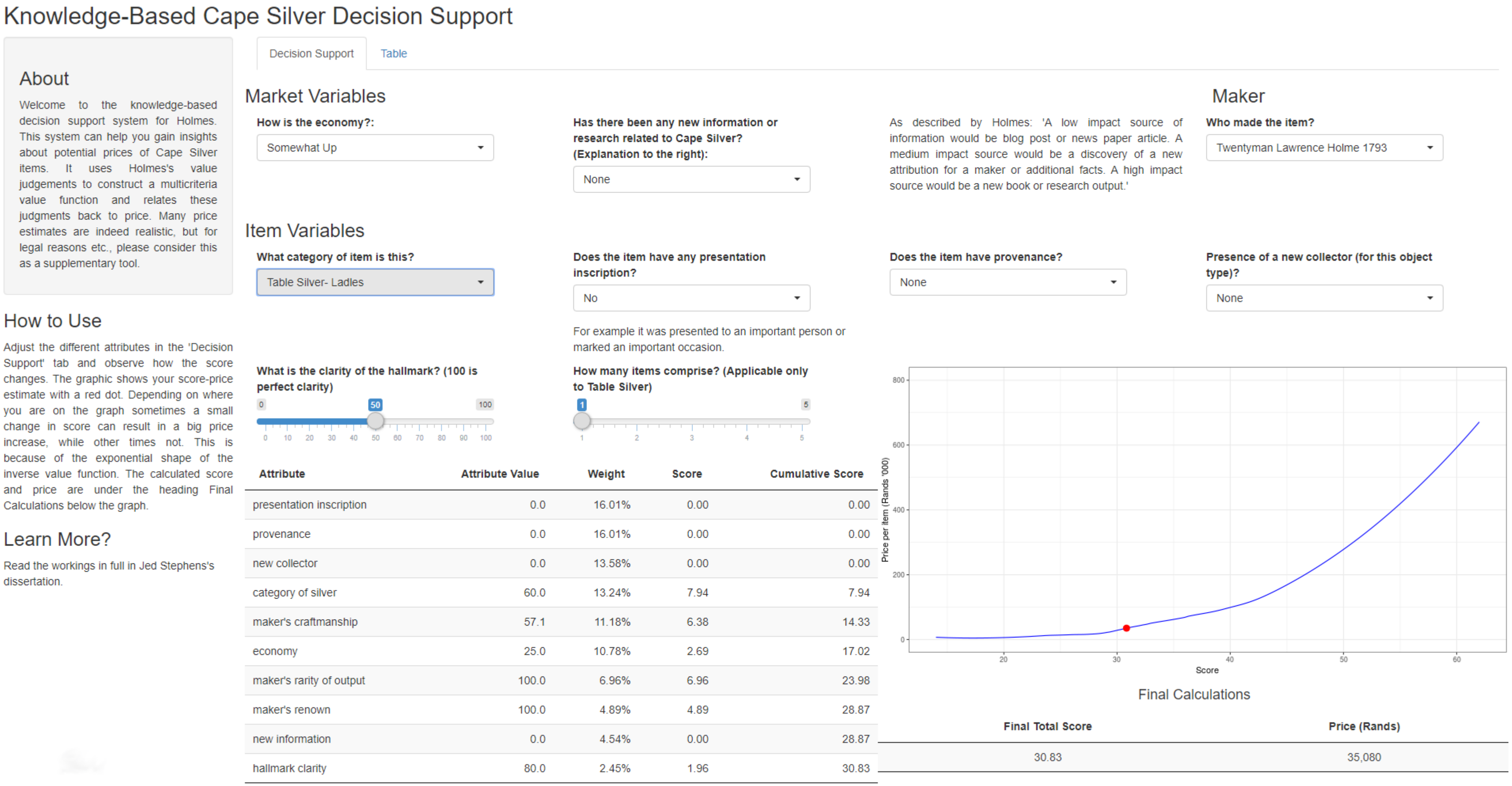


Figure 8.4: Pricing a ladle when the economy is somewhat up.

Chapter 9

Conclusion

This work set out to investigate the use of multicriteria decision modelling to assist in the pricing of non-reproduceable, desirable goods which are describable by their characteristics. For the research to be useable by those who contributed, as well as additional future role-players, it aimed to provide decision support systems as research outputs. At its heart this study has sought to develop the use of multicriteria modelling as a meaningful supplement or alternative to that of the current decision support aids: historical data and price guides. The following is a review of the work.

9.1 Multicriteria Analysis in the Arts

Understanding the role that multicriteria analysis is capable of playing in the Arts, requires highlighting the role that historical data-based decision aids have hitherto played to the field of antique silver (and other art fields). The historical decision aid shows that objects have different attributes (eg. makers or category), but does not explain why or how these differences contribute to the value of an object. The role of multicriteria analysis is to focus on the *why* by constructing a bottom-up understanding of the contributors to an object's price.

Multicriteria analysis can bring detailed numeric information about value judgements. Attribute importance weights provide a quantitative evaluation of hitherto qualitative concepts suggested by respected art resources, such as Pickford (2010); Rosenstein (2009); Waldron (2001). In doing so it presents a new method of expressing this information (i.e. in a value function with an attribute weight), which makes it easier for both DMs and other role-players to speak precisely about the relative importance of different attributes.

However, it should be noted that while multicriteria analysis can play an important role in price decisions, it should not replace historical decision aids – rather these two DSS should be viewed as complementary. Without historical record-based decision aids it is very difficult to calibrate and evaluate a knowledge-based DSS. From a DM's perspective there is little point in creating a knowledge-based decision aid which systematically prices all articles too high. This is the role of the historical data decision aid: to recalibrate the DM's understanding. While historical data captures aggregate trends over time, MC models can help explain why these changes are occurring.

The analyst's experience with DMs has lead to some insights into when multicriteria models are useful. From the perspective of retail-based DMs, multicriteria modelling without price data can still be useful when considering trade-offs between two objects: for example one object is made by maker a and of category f, while the other is made by maker b and of category g. Multicriteria modelling allows the DM to isolate the effects (or 'importance') of these differences in maker or category. Understanding more about trade-offs can play an important role for the collector or academic who wishes to learn more about the field. However, for the retail DM it is only when price is related to their multicriteria model score that it becomes a distinguished decision support tool. The nature of this work focusses on retail pricing. A different pricing structure altogether may have been solicited if the DM were for example a museum curator who would be concerned about the object's value to society.

Multicriteria analysis has the potential to be an extremely beneficial, complementary tool in decision-making processes in the Arts.

9.2 Further Research

To the author's knowledge, this is the first study in the use of multicriteria analysis in the Arts. As such this study has opened up new and exciting avenues for further research. This research takes two forms: refining and resolving. Refining seeks to improve our existing understanding based on the current models. For example, inter-breakpoint (within region) sorting for makers can occur. Currently the different maker's attributes are classed into one of five (carefully chosen) breakpoints. However, makers could be 'sub-sorted' within that breakpoint region to gain a more refined measure. This review process could also suggest that certain makers be moved into a different breakpoint.

Another refinement would be to incorporate attributes such as condition into the final score. The suggestion of Holmes would be to include this as a final multiplicative consideration on the score. To achieve this would require additional data. The approach should also be validated by other DMs in this field.

This work acknowledges limitations in the relationship between MCDM value score and price. As already highlighted there is variability not only inter-DM, but also intra-DM. Such variability impacts on the DSS's ability to determine prices without wide confidence intervals; especially at high prices where the confidence interval could be considerably large. This is also an artefact of the way the MCDM-to-price utility model is constructed: relating all scores with all prices. An alternative as previously discussed would be, for example, to relate low-score items in a separate model.

Most multicriteria models would improve with additional price data. Exploring methods – such as the theory of design of experiments – which simulate this data, could prove to be a very fruitful endeavour. Simulated data also provides one option for tackling the sparseness of high price items exhibited by Holmes. The advantage of simulated data is, for example, that the same maker would appear multiple times across different categories. It is then possible to – at least partially – determine if the resultant score is too low for a particular maker, regardless of the category (pointing to a value judgement on the maker being too low) or for all makers generally in a specific category (pointing to too low a category value). Although this example is given for makers and categories, it applies to all attributes.

Additionally, we could refine the current multicriteria models further by trying to understand the role of inflation. Future research could try to internalise inflation into these models (as was done with Holmes' economy attribute) or understand how best (according to the DM) to adjust for it.

We focus now on how future research can aim at resolving some of the outstanding issues that this dissertation highlights. For example, what is the best method to incorporate mass in instances where this is relevant for the origin (eg. in the case of British silver)? This problem is sidestepped in the current research by the structuring and nature of the problems (Cape origin) tackled. Further, investigations of the utility functions for collectors and a comparison to those of retail DMs could pose interesting future research. As previously discussed, it is expected that the utility (and inverse utility) function for a collector may present a more risk-prone perspective. A further investigation of the algebraic forms of utility functions which may suit the existing risk adverse inverse utility curves would be of interest. The construction of utility functions which incorporate confidence intervals is also an area for future research.

On a different note, construction of criteria for craftmanship is a research interest that would appeal to those interested in researching antique silver objects. Another topic for those with a specialist interest in Cape Silver would be research into the relative output of the Cape silversmiths – thus settling the question about which silversmith's output is the rarest.

More broadly than refining or resolving, future research could look at identifying the type of utility value functions that occur for other non-reproduceable, desirable, describable-by-characteristics objects which exist in the Arts: eg. fine art and jewellery.

9.3 Contributions

This research has had impact on three fields: discrete choice experiments (DCEs), multicriteria decision analysis (MCDA) and antique silver. As a contribution to the DCE literature, this research attempted to integrate the techniques discussed as a method of individual inverse preference elicitation. This study found that DCE experiments do not appear suitable to be used in this way. The appropriateness of using discrete choice to gain insights about a group may nonetheless be insightful for further research. When used in conjunction with individual preference information, group information may be useful in facilitating a group discussion. `ExpertChoice`, the R package developed to construct D-optimal DCEs, provides the first open-source implementation of useful functionality for the discrete choice experiment implementer.

As a contribution to the MCDA literature, this work has been an interesting case study of the complete MCDA process: problem identification, problem structuring, model building, using the model to inform thinking, developing actionable choices, and – finally – reiterating the process when it became apparent that the first MCDA process could be improved upon. It does so in the case of an unusual, but nonetheless interesting problem where the ultimate objective is not a score (of the alternatives), but the score's relation with another variable: in this instance price. Finally, this research illustrates how multicriteria decision analysis needs to be converted into useable decision support systems. It provides two decision support systems for Cape Silver: one historical data-based DSS (`CapeSilversmiths`) and one knowledge-based DSS (`CapeSilverDecisions`).

The work presented in this dissertation also contributes to an understanding about which value elicitation techniques worked well in the field of antique silver – which may be applicable to other fields in the Arts. The MACBETH processes worked excellently in the elicitation of the attribute importance weight. However, the necessary comparisons could become tedious when determining the partial value functions for the attributes. In this instance, bisection methods or a direct sketch of the partial value function were useful. For attributes with large numbers of options (for example, there are more than fifty makers in Cape Silver) the use of predefined breakpoints (or class groups) into which the object was rated, worked very well. Interactive cards, such as those used for the maker's attributes, kept the DMs interested and also provided opportunities to include checks. Further it allows for comparisons between DMs answering on the same makers – which contributes to an understanding of the group dynamics in the field. A final experience-based insight was that, in this study, the DMs preferred a more interactive processes as compared to the DCEs.

A note to analysts looking to continue research with these techniques in antique silver (and more broadly in the arts): prepare for interactions on the basis that it is likely to be a DM's first exposure to a multicriteria modelling session. Come equipped to guide the process with clear expectations of interaction time and general availability to engage. Patience in undertaking these interviews can not be understated: for this research, each DM interview took an excess of five hours and at least two visits to complete. As this is a significant time commitment, one is likely to initially encounter resistance from prospective-DMs. 'Gaining a footing' with key members in the relevant community is a good strategy, as this will allow the analyst access to the specialist community.

Another contribution to the MCDA literature is the practical construction and observation of inverse value functions: the relation of model score to price. The use of the inverse value function suggests a different method for gaining insights into the relationship between model variables. For example, to ask a DM to relate their price to a model score could be challenging. This work has identified and characterised the type of value functions identified as risk-adverse and offered substantive justification for how this could relate to capital.

Finally, this research provides many contributions to the antique silver literature. The importance of structuring a problem within a particular market (eg. Cape) is a clear insight. The group problem structuring maps as well as the individual refinements thereof by certain DMs give a more lucid understanding of this market. It forms a considerable base from which future research can be built. In addition, in the sub-field of Cape Silver where there are limited retail experts, this research captures two different and important key DMs. This research also provides insight into Cape Silver makers which would be of interest to antique silver collectors worldwide and especially to academic projects such as the Antique Dealers Research Project (2020). Minor contributions include the expected risk-adverse relationship between price and score, use of maker's cards and the preferred multicriteria modelling process (viz. use of MACBETH). These are all valuable insights for future researchers working in fields of Decorative Arts. Finally, the core contribution is the construction of the decision support systems (`CapeSilversmiths` and `CapeSilverDecisions`) which will hopefully assist many different role-players in antique silver. In particular it is likely that they will be useful for the DMs to whom – ultimately – this work is dedicated.

References

Aizaki, H. (2012). Basic functions for supporting an implementation of choice experiments in R, *Journal of Statistical Software, Code Snippets* **50**(2): 1–24.
URL: *http://www.jstatsoft.org/v50/c02/*

Antique Dealers Research Project (2020). About the antique dealers research project.
URL: *https://antiquedealers.leeds.ac.uk/about/*

Art Price (2019). Art price. https://www.actusnews.com/fr/ARTPRICE/cp/2017/10/04/artprice-celebrates-its-20th-birthday-editorial-by-thierry-ehrmann-founder-and-ceo-of-artprice.
URL: *https://www.artprice.com/*

Arts & Crafts Exhibition Society (1888). *Catalouge of the First Exhibition MDCCCLXXXVIII*, Chiswick Press.

Bana E Costa, C. A., De Corte, J.-M. and Vansnick, J.-C. (2012). Macbeth, *International Journal of Information Technology & Decision Making* **11**(02): 359–387.
URL: *https://doi.org/10.1142/S0219622012400068*

Banxia Software Ltd. (2019). Decision explorer.
URL: *https://banxia.com/dexplore/*

Belton, V. and Stewart, T. (2002). *Multiple criteria decision analysis: an integrated approach*, Springer Science & Business Media.

Belton, V. and Stewart, T. (2010). Problem structuring and multiple criteria decision analysis, *Trends in multiple criteria decision analysis*, Springer, pp. 209–239.

Burgess, L. and Street, D. J. (2005). Optimal designs for choice experiments with asymmetric attributes, *Journal of Statistical Planning and Inference* **134**(1): 288–301.

Culme, J. (2014). *British Silver Boxes 1640-1840: The Lion Collection*, Antique Collectors' Club.
URL: *https://books.google.co.za/books?id=dpzFoQEACAAJ*

Delieb, E. (2002). *Silver boxes*, 3rd edn, Antique Collectors' Club, Woodbridge, England Wappingers' Falls, NY.

Draper, P., de Ricquebourg, A. D. and Clacher, I. (2018). Auctions, market efficiency, and the trade in second-hand and antique silver, *Economics Letters* **162**: 45 – 48.
URL: *http://www.sciencedirect.com/science/article/pii/S0165176517304391*

Elzhov, T. V., Mullen, K. M., Spiess, A.-N. and Bolker, B. (2016). *minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds*. R package version 1.2-1.
URL: *https://CRAN.R-project.org/package=minpack.lm*

Greco, S., Mousseau, V. and Słowiński, R. (2008). Ordinal regression revisited: Multiple criteria ranking using a set of additive value functions, *European Journal of Operational Research* **191**(2): 416 – 436.
URL: *http://www.sciencedirect.com/science/article/pii/S0377221707008752*

Grömping, U. and Xu, H. (2014). Generalized resolution for orthogonal arrays, *The Annals of Statistics* **42**(3): 918–939.

Heller, D. (1949). *A history of Cape silver : 1700-1870*, D. Heller, Cape Town.

Heller, D. (1953). *Further researches in Cape Silver*, Maskew Miller, Cape Town.

Hensher, D. A., Rose, J. M. and Greene, W. H. (2015). *Applied Choice Analysis*, 2 edn, Cambridge University Press.

Horne, J. (2018). *choiceDes: Design Functions for Choice Studies.* R package version 0.9-3.
URL: *https://CRAN.R-project.org/package=choiceDes*

Kahneman, D. (2011). *Thinking, fast and slow*, Macmillan.

Keeney, R. L. (1996). *Value-focused thinking*, Harvard University Press.

Kuhfeld, W. F. (2010). Marketing Research Methods in SAS Experimental Design, Choice, Conjoint, and Graphical Techniques.

Louviere, J., Islam, T., Wasi, N., Street, D. and Burgess, L. (2008). Designing discrete choice experiments: Do optimal designs come at a price?, *Journal of Consumer Research* **35**(2): 360–375.
URL: *http://www.jstor.org/stable/10.1086/586913*

Louviere, J. J., Street, D., Burgess, L., Wasi, N., Islam, T. and Marley, A. A. J. (2008). Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information, *Journal of Choice Modelling* **1**(1): 128 – 164.
URL: *http://www.sciencedirect.com/science/article/pii/S1755534513700253*

Marttunen, M., Lienert, J. and Belton, V. (2017). Structuring problems for multi-criteria decision analysis in practice: A literature review of method combinations, *European Journal of Operational Research* **263**(1): 1 – 17.
URL: *http://www.sciencedirect.com/science/article/pii/S0377221717303880*

Miller, J. (2017). *Miller's encyclopedia of world silver marks*, Miller's a division of Mitchell Beazley, imprints of Octopus Publishing Group Ltd, London.

Miller, M. and Miller, J. (1979). *Miller's Antiques Price Guide*, M.J.M. Publications.

Morrison, M. N. (1936). *The Silversmiths and goldsmiths of the Cape of Good Hope, 1652-1850*, Mollie N. Morrison, S.l.

Niklewicz, D., Matula, L. and Whetstone, W. (2016). *World Hallmarks (Volume II)*, Hallmark Research Institute, San Francisco, CA.

Palmer, C. E. and Forsyth, C. J. (2006). Antiques, auctions, and action: Interpreting and creating economic value, *The Journal of Popular Culture* **39**(2): 234–259.
URL: *http://onlinelibrary.wiley.com/doi/10.1111/j.1540-5931.2006.00231.x/abstract*

Pickford, I. (2010). *Antique Silver*, reprinted in larger format edn, Antique Collectors' Club, Woodbridge, Suffolk, England.

Pidd, M. (2009). *Tools for Thinking: Modelling in Management Science*, Wiley.
URL: *https://books.google.co.za/books?id=lQDtPAAACAAJ*

Pogrebin, R. (2019). Clean house to survive? museums confront their crowded basements, *New York Times* .
URL: *https://www.nytimes.com/interactive/2019/03/10/arts/museum-art-quiz.html*

Power, D. J. (2002). *Decision support systems: concepts and resources for managers*, Greenwood Publishing Group.

Rose, J. M. and Bliemer, M. C. J. (2009). Constructing efficient stated choice experimental designs, *Transport Reviews* **29**(5): 587–617.
URL: *https://doi.org/10.1080/01441640902827623*

Rosenhead, J. and Mingers, J. (2001). *Rational Analysis for a Problematic World Revisited: Problem Structuring Methods for Complexity, Uncertainty and Conflict*, Wiley.
URL: *https://books.google.co.za/books?id=i30eAQAAIAAJ*

Rosenstein, L. (2009). *Antiques: The History of an Idea*, Cornell University Press.
URL: *http://www.jstor.org/stable/10.7591/j.ctt7zb20*

Saaty, T. L. (1980). The analytic hierarchy process mcgraw-hill, *New York* **324**.

South African Antiques Dealers Association (2019). Directory of antiques dealers.

Stephens, J. (2017). Antique silver as a characteristic good: Aesthetics & price.

Stewart, T. J. (1996). Robustness of additive value function methods in mcdm, *Journal of Multi-Criteria Decision Analysis* **5**(4): 301–309.

Stewart, T. J. (1999). Evaluation and refinement of aspiration-based methods in mcdm, *European Journal of Operational Research* **113**(3): 643 – 652.
URL: *http://www.sciencedirect.com/science/article/pii/S0377221798001428*

Stoller, M. A. (1984). The economics of collectible goods, *Journal of Cultural Economics* **8**(1): 91–104.
URL: *http://www.jstor.org.ezproxy.uct.ac.za/stable/41811143*

Street, D. and Burgess, L. (2007). *The Construction of Optimal Stated Choice Experiments: Theory and Methods*, Wiley Series in Probability and Statistics, Wiley.

Street, D. J., Burgess, L. and Louviere, J. J. (2005). Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments, *International Journal of Research in Marketing* **22**(4): 459 – 470.
URL: *http://www.sciencedirect.com/science/article/pii/S0167811605000510*

The Editors of Encyclopaedia Britannica (2018). The arts.
URL: *https://www.britannica.com/topic/the-arts*

The Editors of Encyclopaedia Britannica (2019). Decorative art.
URL: *https://www.britannica.com/art/decorative-art*

The Silver Society (2019). Silversmiths.
URL: *http://thesilversociety.org/silversmith.php*

Therneau, T. M. (2015). *A Package for Survival Analysis in S*. version 2.38.
URL: *https://CRAN.R-project.org/package=survival*

Therneau, T. M. and Grambsch, P. M. (2000). *Modeling Survival Data: Extending the Cox Model*, Springer, New York.

Traets, F. (2019). *idefix: Efficient Designs for Discrete Choice Experiments*. R package version 0.4.0.
URL: *https://CRAN.R-project.org/package=idefix*

Waldron, P. (2001). *The price guide to antique silver*, Antique Collectors' Club, Woodbridge.

Welz, S. (1976). *Cape Silver & Silversmiths*, A.A. Balkema.

Whetstone, W., Niklewicz, D. and Matula, L. (2010). *World Hallmarks (Volume I)*, Hallmark Research Institute, San Francisco, CA.

Xu, H. and Wu, C. J. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs, *The Annals of Statistics* **29**(4): 1066–1077.

Appendices

Appendix A

Questionnaire

This appendix describes the questionnaire used as part of the problem structuring process. The first questionnaire section, *World View on Value of Antiques*, investigated the DM's thoughts about the use of intuition and systematic thinking in determining a price. It also sought to understand if a DM's estimates are rigid or flexible, and probed for concepts that may be useful to explore more in the formal problem structuring. The second questionnaire section, *The Market for Antique Silver*, requested background about the contemporary antique silver trade. *Decision Aids for Antique Silver* (the third section) asked questions about how DMs engage with the currently available decision support tools (these were described in Chapter 1), as well as whether changing fashions influence the use of these tools. The final questionnaire section, *Pricing Decisions*, aimed to gauge if the DM prices in a currency per gram framework i.e. a premium mark-up on the silver scrap price (see Chapter 3). As previously noted this was historically how objects of antique silver were priced. This questionnaire section also sought to understand if the DM considers himself/herself a *price-setter* or *price-taker* and against which reference point the DM's pricing decisions are formed. An interesting final question is whether the respondent sees himself/herself as adding value through provenance (this term is explained in Table 5.1, page 22). The full questionnaire can be found here.

Section 1

Prefunctory questions

Answers to be captured by the interviewer

Full name	Location of interview	Field of expertise
Full name	Location	antique silver

The institution that the interviewee represents is

choose type

Previous Proceed

Section 2

Interviewee Personal Questions

Answers to be captured by the interviewer

Date of birth

Date of birth. Format: DD/MM/YYYY

A synopsis of your career starting from your current role

Career Synopsis

Your first year of full time work in antique silver

Year. Format: YYYY

Previous Proceed

Section 3

World View on Value of Antiques

Answers to be captured by the interviewee

Q1 Which of the following statements do you agree with most:

- Estimating pricing is an art, not a science.
- Estimating pricing is a science, not an art.
- Estimating pricing is predominately a science with an element of art.
- Estimating pricing is predominantly an art with an element of science.

Q2 When pricing an object of antique silver to what extent do you rely on your intuition (about what the price is)?

highly reliant reliant unreliant highly unreliant

Q3 How would you describe your approach when pricing an object in antique silver?

highly systematic systematic unsystematic highly unsystematic

Q4 Do you believe that the aesthetic merit of an object influences its potential price?

Small Extent Large Extent

Q5 Your price estimates are:

Where between Rigid and Flexible?

Rigid Flexible

Where between Personal and Objective?

Personal Objective

Q6 The price of an object is equivalent to the worth of the object?

strongly disagree disagree neutral agree strongly agree

Q7 Could all the aspects influencing the price of an object be measured?

strongly disagree disagree neutral agree strongly agree

If you think there are aspects that are difficult to measure please note these below:

Kindly seperate the different aspects with commas.

The following table lists some common responses to the above question. Kindly rate these:

Question	Ratings				
Culture	strongly disagree	disagree	neutral	agree	strongly agree
Auction dynamics	strongly disagree	disagree	neutral	agree	strongly agree
Fashions	strongly disagree	disagree	neutral	agree	strongly agree

Q8 A machine would be able to price an object, assuming all the aspects could be measured?

strongly disagree disagree neutral agree strongly agree

Previous Proceed

Section 4

The Market for Antique Silver

Answers to be captured by the interviewer

Q1 **Of your clientele...**

- More than 75% of the buyers are known to me
- Between 50% and 75% of the buyers are known to me
- Between 25% and 50% of the buyers are known to me
- Less than 25% of the buyers are known

Q2 **How well do you know your clientele's preferences?**

- Intimate knowledge of their preferences
- Some knowledge of their preferences
- Limited knowledge their preferences

Q3 **Of your clientele you could identify...**

- More than 75% of the buyers by name
- Between 50% and 75% of the buyers by name
- Between 25% and 50% of the buyers by name
- Less than 25% of the buyers by name

Q4 **How many new buyers join your regular clientele in a given year?**

Number of new regular clientele

Q5 **How many of your regular clientele exit in a given year?**

Number of exiting regular clientele

Q6 **Is it possible to differentiate between buyers who are investors and buyers who are purchasing for the appreciation (love)?**

Yes No

Kindly comment further?

Comments.

(a) What percentage of your clients would you attribute the main reason for purchase as being:

Appreciation (Love) **Investment**

Q7

Appreciation (Love) Proportion | Investment Proportion

(b) What percentage would the average client (of your business) attribute to:

Appreciation (Love) **Investment**

Appreciation (Love) Proportion | Investment Proportion

Q8 **Would a buyer be able to recoup the purchase price of the object a year hence from purchase?**

Description

Q9 **How often do you see objects reappear?**

Description

Q10 **Would it, in your estimation, be viable for a client to purchase an object of antique silver and sell it overseas to recoup at least the purchase price?**

Response

Q11 **Would you say that clients purchase antique silver with the purpose of doing the above, thereby hedging funds overseas?**

Response

Section 5

Decision Aids for Antique Silver

Answers to be captured by the interviewee

Please consider a decision aid in the form of a database giving details of objects in antique silver which have sold previously. There may not be one for antique silver, but as an example, for art there is a well known decision aid database called Art Price (artpriceindex.com). Let us refer to all the records in this decision aid as *historical records*. The following questions relate to the role of historical records in assisting you to decide on a price.

Q1 **Using the following table can you please score the age of historical sale data in assisting you to make a pricing decision**

Years Old	Time Period	Chosen score		
1	Last year	Relevant	Somewhat relevant	Irrelevant
2	Last two years	Relevant	Somewhat relevant	Irrelevant
3	Last three years	Relevant	Somewhat relevant	Irrelevant
5	Last five years	Relevant	Somewhat relevant	Irrelevant
10	Circa 2010	Relevant	Somewhat relevant	Irrelevant
20	Circa 2000	Relevant	Somewhat relevant	Irrelevant
30	Circa 1990	Relevant	Somewhat relevant	Irrelevant
40	Circa 1980	Relevant	Somewhat relevant	Irrelevant
50	Circa 1970	Relevant	Somewhat relevant	Irrelevant
older	Before 1970	Relevant	Somewhat relevant	Irrelevant

Q2 **How far back do you currently go?**

answer in years

Q3 **Do you ever adjust historical data for inflation? (i.e. use real prices?)**

Yes No

Q4 **In the last five years, how many fashions have you observed:**

Started | Ended

Started: answer in years | Ended: answer in years

Can you name these?

Comments.

Q5 **In the ten five years, how many fashions have you observed:**

Started | Ended

Started: answer in years | Ended: answer in years

Can you name these?

Comments.

Q6 **What other information would you like a decision aid to offer? What other information do you think would be useful for a decision aid to offer?**

Description

Previous Proceed

Section 6

Pricing decisions

Answers to be captured by the interviewer

Q1 **Have you ever used the idea of premium/markup on the material cost as a concept to inform your pricing decision?**

Yes No

Any further comment

Comments.

Q2 **When discounting an object does this action affect the true value of the object?**

Yes No

Any further comment

Comments.

Q3 **What would be your rationale for discounting the price of an object from its initial value?**

Q4 **How do you perceive your prices:**

- Less dear than market value
- Market value
- More dear than market value

Q5 **To what extent does your liking of an object influence your estimated price:**

- Not at all influenced
- somewhat influenced
- very influenced

Q6 **Do you tend to underprice objects that you personally dislike?**

Yes No

Any further comment

Comments.

Q7 **Similarly, do you tend to overprice objects that you personally like?**

Yes No

Any further comment

Comments.

Q8 **Do you ever price with a specific buyer in mind?**

Yes No

Any further comment

Comments.

Q9 **With which statement that you would agree with the most:**

- The pre-auction estimate of an object (from a reputable auctioneer) on the whole reflects the market value.
- The pre-auction hammer price estimate (the pre-auction price of an object adjusted to include auctioneer's commission) on the whole reflects the market value an object.
- The winning bid price on the whole reflects the market value of an object.
- The hammer price (winning bid price adjusted to include auctioneer's commission) on the whole reflects the market value of an object.

Please comment on your selected response to Q9

Comments

Q10 **Would objects being bought from your establishment be worth more *because* they were bought from your establishment i.e. do you create value through provenance?**

Yes No

Any further comment

Comments.

Q11 **Estimate the percentage of the object's price that could be attributed to the future provenance that your establishment changes its value.**

answer in percentage points

Previous Proceed

Appendix B

SODA I Cognitive Maps

Figures B.1 and B.2 are a sample of cognitive maps from some of the interviews. Each figure presented is accompanied by an exposition.

For respondent Agatha her modus operandi was to apply a fixed mark-up over cost price. Hence in trying to understand what will be sale price, concept 12, all inputs lead from the purchase price, concept 9. Agatha is in a position that not all items she purchases need to be sold, hence a "sense of preservation" combined with the "right" purchase price may lead Agatha to keep the item for her personal collection. Some of the familiar concepts included in the strategic map are linked with Agatha wanting to purchase the item. One can see age (concept 4), hallmark (clarity and origin were both implied) (concept 6), the design (i.e. quality of the craftmanship for a particular category of item) (concept 2).

An interesting train of Agatha's thoughts surrounded English silver which Agatha did not value highly in and of itself. Rather Agatha considered English silver by its weight and its weight would influence if she would consider a purchase (and naturally at what price). An interest of Agatha's is Scandinavian silver which she thinks is generally undervalued in South Africa.

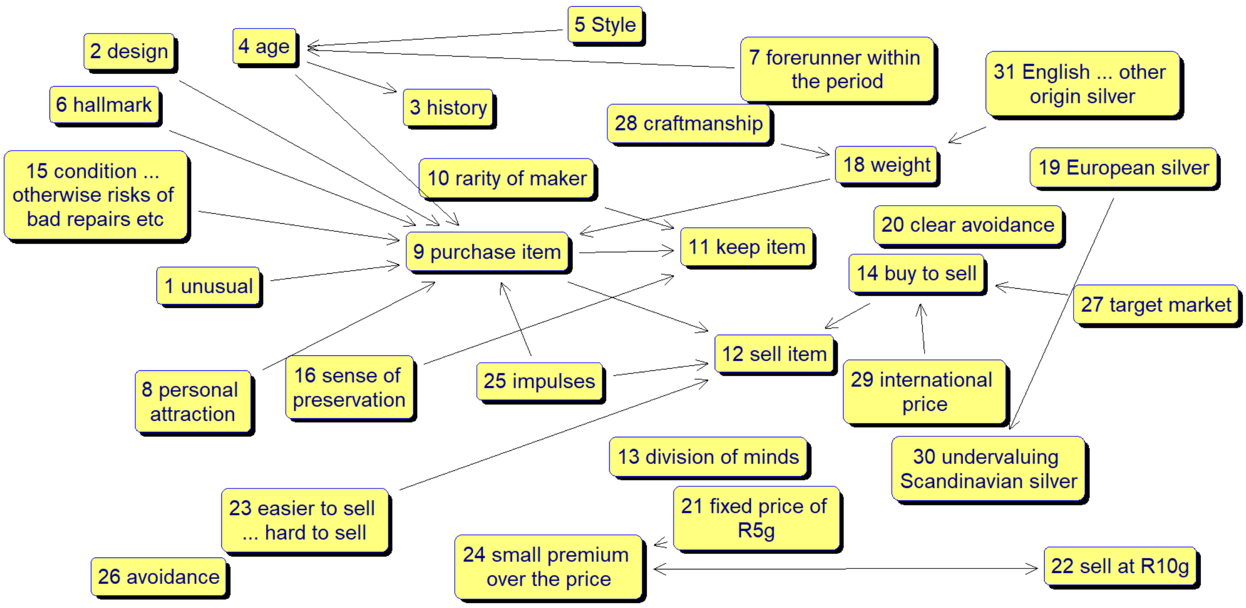


Figure B.1: Agatha's Cognitive Map

For Christie saleability was at the heart of what creates value. If an object could not be sold because it had no saleability then it did not have value. What dictated saleability was demand. What informed demand was the style of the object (concept 8). The style of the object also dictated its beauty (concept 13) which was in turn linked into by many other concepts. In Christie's words: "an ugly thing remains just an ugly thing". This implied that there could be no benefit from an ugly piece of silver. For her silver from the 1900-1950 period which is Art Deco (concept 11) and Art Nouveau (concept 12) were particularly distinguished in terms of their style. Christie only bought (and consequently sold) objects that appealed to her personal taste. The concepts of craftmanship (concept 15), category rarity (concept 5) and origin (concept 17) were all distilled into the strategic map. Additionally the idea of certain periods having a particular style was captured by the age concept in the strategic map.

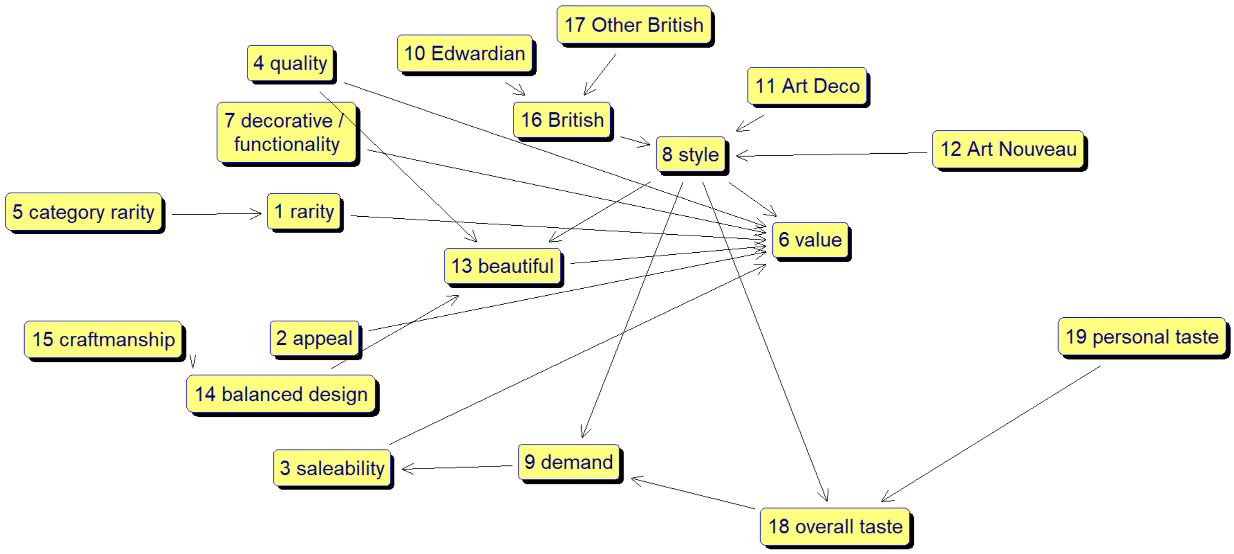


Figure B.2: Christie's Cognitive Map

Appendix C

The `ExpertChoice` R package

This information is current at the time of submission of this dissertation. For the latest changes and updates to the `ExpertChoice` package please see: https://github.com/JedStephens/ExpertChoice

The need for the `ExpertChoice` R package emerged from the methodological desire to implement a discrete choice experiment. There exists a lack of comprehensive open source software to assist in the design of discrete choice experiments. Currently there are three R packages on CRAN that have some overlap with `ExpertChoice`: choiceDes (Horne, 2018), idefix (Traets, 2019) and support.CEs (Aizaki, 2012). Two of these packages are no longer under active development and some of the functions have not been maintained and consequently no longer work. Two packages also lack documentation making it difficult for all but experts in this field to use. `ExpertChoice` provides a unified framework suitable for a first time learner to understand how to design an experiment and convert this experiment into a discrete choice. Its scope is also wider and more current than the above alternate packages.

The following two documents, in this section, are the vignettes that have been written for `ExpertChoice`. As package vignettes the documents can standalone from this dissertation, but were nonetheless, a research output of this dissertation. The audience of these vignettes is obviously broader than this dissertation.

Theoretical introduction to `ExpertChoice` is the first vignette: its objective is to explain the theory of experimental design and discrete choice design. It focusses on explaining how efficiently measure tests play an important role in the designing process. The silver object choice experiment, analysed in this dissertation, is one of the two examples in this vignette. A hypothetical choice experiment on a restaurant is another.

The second vignette, Practical introduction to `ExpertChoice`, provides a worked examples of both experimental designs. The worked examples make it clear how this procedure could be adapted for the reader's own experiment. Some of the more advanced functionality of the package is explored in particular with the restaurant example.

`ExpertChoice` now provides a unified open source alternative to many routines previously only available in SAS and Ngene.

Theoretical introduction to `ExpertChoice`

Jed Stephens

February 10, 2020

1 Purpose

The purpose of this vignette is to provide a theoretical explanation of how to design efficiently. Understanding this in conjunction with the `ExpertChoice` package will allow you to design experiments and discrete choice questionnaires in one paradigm of discrete choice experiments.

Hensher et al. (2015, 287) conclude by explaining that Burgess and Street (2005) and Street and Burgess (2007) launched a literature of optimal stated choice experiments which are based on the multinomial logistic regression and optimal linear experiments. Since then the stated choice/discrete choice literature has expanded in many directions and Hensher et al. (2015) explore this in their comprehensive introduction to Applied Choice Analysis. One large contention that this literature is currently grappling with is the difference between D-optimal and D-efficient (Rose and Bliemer, 2009; Walker et al., 2018). This package, `ExpertChoice`, provides an easy, detailed implementation of the Burgess and Street (2005) literature explaining how to create a DCE in this paradigm. As a paradigm it is gaining renewed interest in particular because it creates experiments that are very robust and are especially suited to situations where there is no prior knowledge.

Why is the term 'experiment' used to describe this process and its literature? Oxford dictionary (2019) gives the following definition for experiment: "A scientific procedure undertaken to make a discovery, test a hypothesis, or demonstrate a known fact." The idea that this definition expresses is that there are variables or attributes which are altered, systematically, and it is the effects of these alterations which is of interest to discover.

2 Designing an experiment

The objective of designing a good experiment is a simple one: ensure that the variables of interest are transparently testable to a chosen satisfactory level. Designing an experiment starts before becoming wrapped up in theoretical considerations as to how best to achieve this. The important first step is defining the relevant attributes and their levels. Doing this requires some prior insight. Problem structuring methods (such as those described in Belton and Stewart (2002)) can be incorporated with the first two design stages described by Hensher et al. (2015, 194-201). In the linked Practical Introduction to `ExpertChoice` this process is described as Step 0.

The next section will introduce two examples of experiments where this groundwork has already been done. The first, a restaurant experiment, is relatively of a small size, but has interesting considerations. The second, a silver object experiment, is larger and was the motivation for this package being written. Both examples are fully worked in the linked Practical Introduction to `ExpertChoice`.

2.1 Describing your experiment

An experiment starts by describing the variables which change. It is Step 0. The variables, also referred to as attributes, can be ordered (ordinal) or unordered (categorical). This distinction is important if the intention is to convert the experiment into a discrete choice experiment, but not so much otherwise. This section will describe two experiments: one with all variables unordered (categorical) and another with all variables ordered. It is possible to also have an experiment with a mixture of the two.

2.1.1 Unordered Variable Experiment: Restaurant Experiment

Imagine you own a restaurant that only serves a set menu. As the menu is set your patrons never choose what they are getting for each part of the menu or infact what they are getting on the day. You want to experiment with different set menus to see not only what meals patrons enjoy, but also which starter, main and dessert combinations work well together. In your repertoire you have the following recipes[1]:

starter = {Tomato Soup, Duck Rillettes, Seafood Chowder}
main = {Roast Pheasant, Pan Fried Hake, Pork Belly, Mushroom Risotto, Sirloin Steak, Vegetable Bake}
dessert = {Sticky Toffee Pudding, Chocolate & Hazelnut Brownie, Cheesecake}

Table 1 formally describes this experiment. The different levels of the variables (z_l) are labelled starting from 1 upwards. Starting at 1 is the convention for unordered variables. This design would be called a $3^2 6^1$ design because there are two variables each with three levels and one variable with six levels. As a result there would be 54[2] different variations of set menus that could be offered.

z	**attribute/variable name** (z)	z_l	**level name** (z_l)
1	Starter	1	Tomato Soup
1	Starter	2	Duck Rillettes
1	Starter	3	Seafood Chowder
2	Main	1	Roast Pheasant
2	Main	2	Pan Fried Hake
2	Main	3	Pork Belly
2	Main	4	Mushroom Risotto
2	Main	5	Sirloin Steak
2	Main	6	Vegetable Bake
3	Dessert	1	Sticky Toffee Pudding
3	Dessert	2	Chocolate & Hazelnut Brownie
3	Dessert	3	Cheesecake

Table 1: Design for research on different set menus. (A $3^2 6^1$ design.)

2.1.2 Ordered Variable Experiment

The silver experiment was an attempt to determine utility functions for the different variables from experts in antique silver. Notice here that unlike in Table 1 the levels of the variables (z_l) are labelled

[1](This menu is adapted from the restaurant 101 Talbot http://www.101talbot.ie/menus/)
[2]This number should not be too suprisingly it comes from the design notation that is $54 = 3^2 \times 6^1$

starting from 0 upwards. This is to make explicit the fact that the $z_l = 0$ level will be used as the base level. In unordered variables the base level can be chosen arbitrarily. This experiment would be described as 5^5 and hence there are 3125 possible combinations.

z	**attribute name** (z)	z_l	**level name** (z_l)	**Description**
1	Makers Renown	0	bottom 50% of makers	common
1	Makers Renown	1	50% to 65% of makers	known to specialists
1	Makers Renown	2	65% to 80%	recognised
1	Makers Renown	3	80% to 90%	famous
1	Makers Renown	4	top 10%	celebrated
2	Technical Perfection	0	below 50% of craftmanship	below average
2	Technical Perfection	1	50% to 65%	good
2	Technical Perfection	2	65% to 80%	meritorious
2	Technical Perfection	3	80% to 90%	distinguished
2	Technical Perfection	4	top 10%	exquisite
3	Category Rarity	0	bottom 20%	common
3	Category Rarity	1	20% to 40%	category rarity
3	Category Rarity	2	40% to 60%	rare
3	Category Rarity	3	60% to 80%	very rare
3	Category Rarity	4	top 20%	exceptional
4	Size (of object)	0	under 125g	petite
4	Size (of object)	1	between 126g and 275g	small
4	Size (of object)	2	between 276g and 600g	medium
4	Size (of object)	3	between 601g and 1200g	large
4	Size (of object)	4	exceeds 1200g	extra large
5	Age (of object)	0	1951-present	
5	Age (of object)	1	1900-1950	
5	Age (of object)	2	1851-1899	
5	Age (of object)	3	1801-1850	
5	Age (of object)	4	before 1800	

Table 2: Design for research on antique silver objects to be answered by experts. (A 5^5 design.)

3 The Factorial Designs

3.1 The Full Factorial Design

The full factorial design is constructed from z attributes each with l levels denoted as z_l. It contains all possible combinations of the levels of the attributes and each row is unique. The full factorial for any design has the unique number of combinations (hence the number of rows of the full factorial) given by the design description. Recall for the restaurant and silver experiments that was $3^2 6^1$ and 5^5 respectively. Table 5, at the end of this document gives the full factorial design for the restaurant experiment.

3.2 The Fractional Factorial Design

It is clear that the full factorial design can quickly become overwhelming if the intention was to implement an experiment where each scenario (i.e. row) as conducted. It would also be onerous, costly and more so depending on what is desired unnecessary. The aim of all experimental design is to get to the results faster and as effectively as possible. Therefore for all but toy examples the researcher must concern themselves with how best to select from the full factorial design to form the fractional factorial design. The fractional factorial design is always (by definition) be contained in the full factorial design.

In general there are three methods to select from the full factorial: column based methods (typically methods originating with Federov), row based methods (mixed integer programming) and the construction of orthogonal arrays (Grömping, 2018; Kuhfeld, 2010). There are merits to each method, but before an extensive discussion can be had it is necessary to explain further some of the efficacy measures for factorial designs. These efficacy measures are meant to guide the selection process.

3.2.1 Efficacy Measures for Factorial Designs

A design's main effects are the effects of the each of the attributes measured at each of the levels. Two-attribute interactions are the interaction effects between z^a and z^b where $a \neq b$. In order to have a two-attribute interaction effect there must be at least two factors. Similarly for three factors. The number of estimable n-attribute interactions are related to a efficacy measure described by Xu and Wu (2001) and Grömping and Xu (2014) as generalised word lengths. That is the nth world length is the n-attribute interactions that the full factorial would support. For the menu example there are 3 attributes hence 3-attribute interactions as such the there are generalised word lengths 0,1,2 and 3. For the silver object design there are 5 attributes hence 5-attribute interactions as such the there are generalised word lengths 0,1,2,3,4,5 and 6. Only the full factorial design is capable of supporting the nth world length. But in many instances one's interest is only in the main effects and/or possibly two level interactions. This comes with the major advantage that a fraction of the full factorial may now be used.

Generalised word lengths are a powerful method of assessing design efficacy which has a strong relationship to two more familiar concepts from the DoE literature: resolution and strength. Strength[3] s is equal to the resolution[4] (r) less 1. The first generalised world length, always the zero word length is always 1 i.e. generalised word length (0) = (1). Thereafter the number of zero length words is the strength of the design. For example in the menu design there are 3 word lengths (as explained previously). For the silver experiment if generalised world length (1) = 0, generalised word length (2) = 0 and generalised word length (3) = 0, but generalised word length (4) = 25, i.e. 3 zero length words, then the strength of the design is 3, hence resolution four.

This is significant because, following Kuhfeld (2010), stated generally if resolution (r) is odd then the effects of order $e = (r - 1)/2$ or less are estimable free of each other. However at least some of the effects of order e are confounded with interactions of order $e + 1$. If r is even then effects of order $e = (r - 2)/2$ are estimable free of each other and are also free of interactions of order $e + 1$. Table 3 gives some commonly chosen designs and an interpreted description.

[3]Strength is traditionally denoted with a number: 1,2,3,4

[4]Resolution is traditionally denoted in roman numerals or in words

resolution	strength	number of zero length words	description
III	2	2	all main effects are estimable free of each other, but some are confounded with two-attribute interactions
IV	3	3	all main effects are estimable free of each other and free of all two-factor interactions, but some two-attribute interactions are confounded with other two-attribute interactions
V	4	4	all main effects and two-factor interactions are estimable free of each other

Table 3: Commonly chosen designs and their efficacy

The full factorial design has the maximum achievable resolution, strength and number of zero length words. Hence, although it is only possible in toy examples, it is best possible design. It also has two other desirable properties: orthogonality and level balance.

The efficacy of a design for a particular specification can be calculated for a specific stipulation. Let $\mathbf{X}$ be the design matrix of the proposed design (typically this is the fractional factorial design) with an intercept and its attributes expanded using standardised orthogonal contrast coding[5]. The information matrix (familiar from theory of the linear model) is $\mathbf{X^TX}$. The number of rows in the proposed design is denoted N_D. The number of rows (or columns) in the symmetric information matrix ($\mathbf{X^TX}$) is denoted as p. The A-efficiency is defined as

$$\frac{100}{N_D} \times \frac{1}{\text{trace}((\mathbf{X^TX})^{-1})/p} \tag{1}$$

and the D-efficiency[6] as

$$\frac{100}{N_D} \times \frac{1}{\det((\mathbf{X^TX})^{-1})^{(1/p)}}. \tag{2}$$

It cannot be overemphasised that A-efficiency and D-efficiency of a design is specific to the particular model matrix expansion of the proposed design. For example assume that a suitable fractional factorial design for the silver objects experiment is given by the matrix $\mathbf{B}$. The design expansion of the matrix $\mathbf{B}$ would be different when estimating only the main effects (viz. Makers Renown, Technical Perfection, Category Rarity, Size and Age) as it would be when an expansion that included some (or all) interactions (viz. Makers Renown, Technical Perfection, Makers Renown $\times$ Technical Perfection, Category Rarity, Size). Let us assume that matrix $\mathbf{B}$ is of resolution IV then, by definition, the A- and D- efficiency of the main effects design will be 100% i.e. fully efficient. Yet, the second proposed expansion of $\mathbf{B}$ (the one including interactions) may or may not be 100% efficient. It will depends on whether those particular interactions ailise each-other. In general to inspect which effects ailise which other expand the $\mathbf{X}$ up to the value of e (see definition earlier) and investigate the information matrix ($\mathbf{X^TX}$).

To make this discussion about model expansions more concrete, Step 5 of the menu experiment, demonstrates how to programme these tests. The following table, Table 4, summarises the results for the different formulations. The notation '+' indicates the variable is added linearly, while the '$\times$' indicates that the variables and its interactions are added. A 36 run orthogonal array with generalised

[5] Any coding can be used in analysis of the completed experiment. The standardised orthogonal contrast coding has attractive properties when designing an experiment as its efficiency measures are normalised to 100% in the case of the optimal design.

[6] D-efficiency is, in general, a relationship between $[\det(C)/\det(C_{\text{optimal}})]$ where C is the information matrix in the case of the linear model viz. $\mathbf{X^TX}$. For linear models the $\det(C_{\text{optimal}})$ is well known.

word lengths $(0) = 1, (0) = 0, (1) = 0, (2) = 0, (3) = 0.5$ was used. This design hence has strength of 2 (the number of zero lengths words is $(3) - (1) = 2$).

Design Expansion	**minimum strength for full efficiency**	**A-efficiency (%)**	**D-efficiency (%)**
starter + main + dessert	2	100	100
starter × main + dessert	3	93.75	97.164
starter + main × dessert	3	93.75	97.164
starter × dessert + main	3	91.304	95.975
starter × main× dessert[7]	4	NA	NA

Table 4:

The information matrix $(\mathbf{X^T X})$ is also telling about the balance and orthogonality of the design. A design is orthogonal when the sub-matrix of $(\mathbf{X^T X})^{-1}$ (excluding the row and column for the intercept) is diagonal. (There may be off-diagonal non-zeros for the intercept.) A design is balanced when all off diagonal elements in the intercept row and column are zero. When a design is both simultaneously balanced and orthogonal, the $(\mathbf{X^T X})^{-1}$ matrix is diagonal and $(\mathbf{X^T X})^{-1}$ is equal to $\frac{1}{N_D}\mathbf{I_{(p\times p)}}$ (Kuhfeld, 2010, 63). Such a design is a 100% efficient design. That is, practically speaking, the design does not in any way influence the results – it has no systematic bias. All designs less than 100% efficient may have balance or orthogonality or neither.

In the conjoint literature there existed a historical preference for designs that are orthogonal, despite the fact that some of these designs may have been very imbalanced and hence rather inefficient. This literature has now migrated to choosing designed based on D-efficiency (Kuhfeld, 2010). (Which may result in designs which are neither orthogonal nor balanced, but are relatively orthogonal and balanced.) In the discrete choice literature there exists a strong preference for a balanced design (Hensher et al., 2015).

4 Experiments without blocks

A major technique in the design of both conjoint and discrete choice experiments is blocking. Blocking is a design of experiment (DoE) term used to describe a situation where different respondents answer different portions of the chosen design. (In more technical terms blocking is the division of the chosen fractional factorial design.) Blocking is a systematic technique of division – typically blocks are mutually exclusive of one-another (so called "no-overlap designs"), but increasingly often with purposeful overlap (so called "minimal overlap designs").

Why block? Sometimes a fractional factorial design may be too large that it can be reasonably answered by one respondent. Blocking breaks the experiment into smaller "bites" for respondents.

The appropriateness of blocking has come under strong theoretical scrutiny (Hensher et al., 2015; Rose and Bliemer, 2009). Their arguments can be summarised as this, typically in large respondent surveys when blocking is used the result can be imbalance of administration of the blocks. Let us assume that a design is separated into four blocks (A, B, C, D). The study has 50 participants. Firstly, four does not divide 50 equally so the researcher must make the choice of which of the blocks to give to the 49th and 50th respondent. Secondly many things could foul a response: the respondent may wish to withdraw from the study, they may be missing questions or have answered illegibly, etc. Let us assume that there are 47 usable responses and that these consist of 12 A blocks, 7 B blocks, 11 C blocks and 17 D blocks which collectively sum to 47. The problem is now self-evident: analysis happens based on the original chosen design. The blocked reconstruction of the original chosen design is fatally flawed it will introduce imbalances (where they never existed) and will struggle to estimate with the same efficacy. This is strong motivation to avoid blocked designs.

4.1 Selecting from the Full Factorial Design

Constructing the fractional factorial design is very much an iterative process of using a selection method evaluating the design and then typically reiterating. Step 3 in the Practical Introduction to `ExpertChoice` demonstrate how to do so using a column based, row based and orthogonal array approach. Hensher et al. (2015) provide a good introduction to these different methods.

5 Moving from Factorial Design to Discrete Choice Design

A discrete choice experiment consists of several choice sets (denoted as the number of runs) with each choice set containing two or more options (denoted as alternatives). The most apparent difference, the fact that in discrete choice there must be choice within choice sets soon takes forefront concern in converting from a fractional design to a discrete choice design.

For now the focus is to review the different techniques currently available to convert from a factorial design to a discrete choice design. The following section is quashed in a warning about efficiency measures for discrete choice experiments. Methods for laying out DCE are the same regardless of the paradigm for evaluating their end results.

- Modulo Methods (these are proposed by Street and Burgess (2007) and have many advantages)
- L^{MA}[8]
- Rotation Method[9]
- mix-and-match Method [10]

5.1 Optimality Measures for MNL Discrete Choice Designs

The efficacy measures for discrete choice designs inherit much of the literature from the fractional designs (used in conjoint analysis). This can be incorporated into a measure of efficacy for DCE that are used to estimate the main effects or the main effects plus two-factor interactions. Optimal designs will, when using D-optimally criterion have the maximum determinant of the Fisher information matrix. For DCE the information matrix is defined to be: $C = B\Lambda B'$. The B matrix is the matrix of contracts for the effects that are to be estimated. Although theoretically possible, currently the `ExpertChoice` package only supports the construction of the B matrix for main effects. Hence the D-optimallity calculated

[8] "The L^{MA} method directly creates a choice experiment design from an orthogonal main-effect array (Johnson et al. 2007). In this method, an orthogonal main-effect array with M times A columns of L level factors is used to create each choice set that contains M alternatives of A attributes with L levels. Each row of the array corresponds to the alternatives of a choice set." (Aizaki, 2012)

[9] "The rotation method uses an orthogonal main-effect array as the first alternative in each choice set; this method creates one or more additional alternative(s) by adding a constant to each attribute level of the first alternative; the kth(¿= 2) alternative in the jth (= 1, 2, ..., J) choice set is created by adding one to each of the m attributes in the k - 1 th alternative in the jth choice set. If the level of the attribute in the k - 1 th alternative is maximum, then the level of the attribute in the k th alternative is assigned the minimum value." (Aizaki, 2012)

[10] "The mix-and-match method modifies the rotation method by introducing the randomizing process. After placing a set of N alternatives created from the orthogonal main-effect array into an urn, one or more additional set(s) of N alternatives are created using the rotation method and placed into different urn(s). A choice set is generated by selecting one alternative from each urn at random. This selection process is repeated, without replacement, until all the alternatives are assigned to N choice sets. These N choice sets correspond to a choice experiment design." Aizaki (2012)

by `ExpertChoice` will, at this stage, always be for main effects only. This may sound limiting, but it will become apparent that achieving a D-optimal DCE for main effects is sufficiently challenging not to warrant further complication. The Λ matrix is the matrix of second derivatives of the likelihood function which "under the null hypothesis of no differences between the effects of the levels of each attribute turns out that Λ contains the proportions of choice sets in which pairs of profiles appear together" (Street and Burgess, 2007, 462). The entries in Λ can be evaluated by counting the occurrences of pairs of profiles and diving by m^2N where N is the number of choice sets (Street and Burgess, 2007, 462). The D-efficency of the design of any design can be given by $[\det(C)/\det(C_{\text{optimal}})]$.

There exists a theoretical $\det(C_{\text{optimal}})$ for main effects only, first determined by Burgess and Street (2005), for the multinomal logit model. The mathematics of this are given most plainly in Street and Burgess (2007). The `ExpertChoice` package has this functionally built into it. See the Practical Introduction, step 9.

Since then "Bliemer and Rose (2014) were able to show that Street and Burgess designs are simply a special case of the more general methods used by other researchers" (Hensher et al., 2015, 310). In particular this more general method requires a larger introduction than is possible here. For the most neutral introduction see Walker et al. (2018). Much of the diffcult boils down to: "D-optimal designs attempt to maximize attribute level differences whereas D-efficient designs attempt to minimize the elements that are likely to be contained within the AVC matrices of models estimated from data collected using the design." (Rose and Bliemer, 2009)

starter	main	dessert
1	1	1
2	1	1
3	1	1
1	2	1
2	2	1
3	2	1
1	3	1
2	3	1
3	3	1
1	4	1
2	4	1
3	4	1
1	5	1
2	5	1
3	5	1
1	6	1
2	6	1
3	6	1
1	1	2
2	1	2
3	1	2
1	2	2
2	2	2
3	2	2
1	3	2
2	3	2
3	3	2
1	4	2
2	4	2
3	4	2
1	5	2
2	5	2
3	5	2
1	6	2
2	6	2
3	6	2
1	1	3
2	1	3
3	1	3
1	2	3
2	2	3
3	2	3
1	3	3
2	3	3
3	3	3
1	4	3
2	4	3
3	4	3
1	5	3
2	5	3
3	5	3
1	6	3
2	6	3
3	6	3

Table 5: The full factorial design for the restaurant experiment. (A $3^2 6^1$ design.)

References

Hideo Aizaki. Basic functions for supporting an implementation of choice experiments in R. *Journal of Statistical Software, Code Snippets*, 50(2):1–24, 2012. URL `http://www.jstatsoft.org/v50/c02/`.

Valerie Belton and Theodor Stewart. *Multiple criteria decision analysis: an integrated approach*. Springer Science & Business Media, 2002.

Leonie Burgess and Deborah J Street. Optimal designs for choice experiments with asymmetric attributes. *Journal of Statistical Planning and Inference*, 134(1):288–301, 2005.

Ulrike Grömping. R package doe.base for factorial experiments. *Journal of Statistical Software, Articles*, 85(5):1–41, 2018. ISSN 1548-7660. doi: 10.18637/jss.v085.i05. URL `https://www.jstatsoft.org/v085/i05`.

Ulrike Grömping and Hongquan Xu. Generalized resolution for orthogonal arrays. *The Annals of Statistics*, 42(3):918–939, 2014.

David A. Hensher, John M. Rose, and William H. Greene. *Applied Choice Analysis*. Cambridge University Press, 2 edition, 2015. doi: 10.1017/CBO9781316136232.

Warren F. Kuhfeld. Marketing Research Methods in SAS Experimental Design, Choice, Conjoint, and Graphical Techniques, 2010.

John M. Rose and Michiel C. J. Bliemer. Constructing efficient stated choice experimental designs. *Transport Reviews*, 29(5):587–617, 2009. doi: 10.1080/01441640902827623. URL `https://doi.org/10.1080/01441640902827623`.

D.J. Street and L. Burgess. *The Construction of Optimal Stated Choice Experiments: Theory and Methods*. Wiley Series in Probability and Statistics. Wiley, 2007. ISBN 9780470053324.

Joan L Walker, Yanqiao Wang, Mikkel Thorhauge, and Moshe Ben-Akiva. D-efficient or deficient? a robustness analysis of stated choice experimental designs. *Theory and Decision*, 84(2):215–238, 2018.

Hongquan Xu and CF Jeff Wu. Generalized minimum aberration for asymmetrical fractional factorial designs. *The Annals of Statistics*, 29(4):1066–1077, 2001.

Practical Introduction to ExpertChoice

```
library(ExpertChoice)
```

The purpose of this example is to present a practical worked example of how to design a conjoint and discrete choice experiment without blocking.

Step 0: Decide on what to test

The process of choosing a design often involes interating over steps 0 to 4. Some designs are more difficult to create than others. The Theoretical Introduction to ExpertChoice presents two designs which are illustrated in this practical vingette.

Here are some practical suggestions as to what makes a good design. 1. In general avoid attributes with only two levels. The design such as the one below suffers because it is difficult to convert from the fractional fractorial of this design into an effient choice experiment design. The lack of efficiency is not from the methods of converting, but inherit in the fact that achieving the minimal overlap when there is only two levels is diffciult.

Step 1: Construct the full factorial

First load the the `ExpertDesign` package into your R environemnent.

```
library(ExpertChoice)
```

Create a list object which specifies the name of the variables as well as thier respective levels.
NB: The levels should be integer sequential and start from 0. The level at 0 is always the reference point. As the proposed design above is a 5^5 I have chosen to denote the object as `attr55`:

```
attri55 <- list(
  maker = c("0", "1", "2", "3", "4"),
  technical = c("0", "1", "2", "3", "4"),
  category_rarity = c("0", "1", "2", "3", "4"),
  size = c("0", "1", "2", "3", "4"),
  age = c("0", "1", "2", "3", "4")
)
```

Calling the list object something like this is advantagous because you could have multiple competing designs still at this stage. The following design is $4^4 2^1$ and then denoted here as `attri4521`:

```
attri4521 <- list(
  maker = c("0", "1", "2", "3"),
  technical = c("0", "1", "2", "3"),
  category_rarity = c("0", "1", "2", "3"),
  size = c("0", "1", "2", "3"),
  age = c("0", "1", "2", "3"),
  provenance = c("0", "1")
)
```

Create the full factorial object. Using the design specification as a suffix remains a handy way of keeping track of the design.

```
ff55 <- full_factorial(attri55)
```

The full factorial will contain many rows. The first five rows and the last five are given below:

```
rbind(head(ff55, 5), tail(ff55, 5))
#>      maker technical category_rarity size age
#> 1        0         0               0    0   0
#> 2        1         0               0    0   0
#> 3        2         0               0    0   0
#> 4        3         0               0    0   0
#> 5        4         0               0    0   0
#> 3121     0         4               4    4   4
#> 3122     1         4               4    4   4
#> 3123     2         4               4    4   4
#> 3124     3         4               4    4   4
#> 3125     4         4               4    4   4
```

Every variable in the full factorial has the standaridised orthogonal contrast applied. These constrasts are very useful when eveluating the efficacy of a design. This is simply illustrative of what the contrasts look like for one of the variables:

```
contrasts(ff55$maker)
#>        [,1]       [,2]       [,3] [,4]
#> 0  1.581139 -0.9128709 -0.6454972 -0.5
#> 1  0.000000  1.8257419 -0.6454972 -0.5
#> 2  0.000000  0.0000000  1.9364917 -0.5
#> 3  0.000000  0.0000000  0.0000000  2.0
#> 4 -1.581139 -0.9128709 -0.6454972 -0.5
```

Step 2: Augment the full factorial

Once the full facotrial is constructed it is possible to augment it with additional information. Many of these augmentations happen as attributes. This includes adding the B-matrix for main effects, an important matrix in DCE efficiency, as described by Street et al… The prefix `aff` is used to refer to the augmented (full) factorial. (You could of course name the object whatever you prefer.)

```
aff55 <- augment_levels(ff55)
#> [1] "Applying B mat"
```

A console log will appear stating that the processes of applying the B-matrix has started. If you do not get this message then the B-matrix cannot be added. (Please open a GitHub issue if this is the case. I am not aware of instances where this should happen.) The B-matrix plays an important role in the choice efficiency of design. Below are ten random rows drawn from the agumented full factorial. Notice the additional of the `levels` column.

```
aff55[sample(nrow(aff55), 10), ]
#>      maker technical category_rarity size age levels
#> 1115     4         2               4    3   1  42431
#> 727      1         0               4    0   1  10401
#> 3033     2         1               1    4   4  21144
#> 3074     3         4               2    4   4  34244
#> 2383     2         1               0    4   3  21043
#> 389      3         2               0    3   0  32030
```

```
#> 2802     1          0                2    2    4   10224
#> 2260     4          1                0    3    3   41033
#> 1714     3          2                3    3    2   32332
#> 358      2          1                4    2    0   21420
```

Step 3: Creating a fractional factorial design.

```
library(AlgDesign)
library(DoE.base)
# library(DoE.MIParray)
```

There are many ways to create a fractional factorial design. See Section … of the associated note Designing conjoint scaling and discrete choice experiments for small sample expert surveys by Jed Stephens for a full discussion. Practically speaking though two methods are designed to be flawlessly integreated into this package. These are the construction of a fractional factorial design using an orthogonal array with either the `DoE.MIParray` or `DoE.base` packages or using D-optimal fractional factorial designs from the `AlgDesign` package.

Ortogonal Arrays (`DoE.MIParray` or `DoE.base`)

Determine feasiability

The function `oa_feasible()` from the `DoE.base` package (`DoE.base::oa_feasible()`) provides many methods for determining if a particular design can be construed with N_D rows. For the silver expert it was found that the following design was feasiable. It is possible to specifiy higher resolution designs. These are always advantageous. See Section … of the associated note Designing conjoint scaling and discrete choice experiments for small sample expert surveys by Jed Stephens for a full discussion.

```
# Design: DF: 17, 32 OA (Resolution II), 64 OA (Resolution III)
nlevels <- unlist(purrr::map(ff55, function(x){length(levels(x))}))
oa_feasible(25, nlevels, strength = 2)
#> no violation of necessary criteria  for strength  2  was found
#> [1] TRUE
```

Using the `DoE.base` package it is possible to construct a 25 .

```
fractional_factorial_55_25 <- oa.design(nlevels = nlevels, columns = "min34")
```

When consturcting a design using the `DoE.MIParray`… The function `mosek_MIParray()` was used to construct the example 64 run orthogonal array included with this package.

```
# Not run because it requires time as well as some setting up if this is your first time.
# See DoE.MIParray for more detials.
#fractional_factorial_55_25 <- gurobi_MIParray(25, nlevels)
```

Note how the functions in DoE use a slightly different notation to index the factors. The base is given level 1 in this package. This is not really a concern because it will be seemlessly converted shortly.

```
head(fractional_factorial_55_25, 10)
#>    A B C D E
#> 1  3 4 5 1 2
#> 2  5 2 4 1 3
#> 3  5 3 1 4 2
#> 4  1 5 4 3 2
#> 5  2 2 2 2 2
#> 6  1 2 3 4 5
#> 7  5 1 2 3 4
#> 8  2 3 4 5 1
#> 9  2 5 3 1 4
#> 10 1 4 2 5 3
```

D-efficient

Not yet discussed or though it should be achieveable with mininal effort. If a reader wishes for this example to be completed before I have done so please open a GitHub issue and I shall happily oblige completing.

Step 4: Searching the full factorial for the chosen fractional factorial design

The ability to use multiple different packages to construct the fractional factorial design is ensured by this step. There can exist small differences between the different methods which require some fiddiling.

The results of the `mosek_MIParray` function are orthogonal arrays without colnames. Hence in this instance the colnames need to be added. This design clearly needed to be made with the full factiorial in mind. Hence the colnames from the `ff4521` object are appropriate. **Note: the colnames from the aff4521 would include the levels column – hence avoid these…**

```
colnames(fractional_factorial_55_25) <- colnames(ff55)
fractional_f55_25 <- search_design(ff55, fractional_factorial_55_25)
```

The result is a fracitonal fractional design. Importantly though the fractional factorial design retains and inherits information from the full factorial such as the standarised orthgonal coding. To mark that many such attributes are held a special attribute is assigned to the object.

```
# Check to see if the searched attribute exists on the fractional_f4521_64 object.
attributes(fractional_f55_25)$searched
#> [1] TRUE
```

Once an object is search converted it is now easy to run diagonsitcs.

Step 5: Determining the efficacy of (full or fractional) factorial designs

The theoertical discussion of these diagonstics is presented extensively in the associated note Designing conjoint scaling and discrete choice experiments for small sample expert surveys by Jed Stephens. See Section…

The generalised world length patterns gives a good overall summary of the design.

```
DoE.base::GWLP(fractional_f55_25)
#>  0  1  2  3  4  5
#>  1  0  0 40 40 44
```

From this we can tell that this fractional factorial design is resolution IV i.e. strength of 3. Hence the all main effects are estimable free of each other, but some are confounded with two-attribute interactions.

The function `fractional_factorial_efficiency` provides a formula based method of investigating the proposed fractional factorial design in more detials. This function also includes in its list of results the GWLP so there is no need to specifiy it.

Two examples are given which follow the two examples in the associated note.

```
# Test for main effects
main_effects <- fractional_factorial_efficiency(~ maker + technical + category_rarity + size +
         age, fractional_f55_25)
#> Your fractional factorial design has an A-efficiency of 100 %
#>  Your fractional factorial design has a D-efficiency of 100 %
```

The resultant object has the following objects:

```
names(main_effects)
#> [1] "X"                     "information_mat"      "inv_information_mat"
#> [4] "lamda_mat"             "inv_diag"             "GWLP"
#> [7] "A_eff"                 "D_eff"
```

Check the package help file for the `fractional_factorial_efficiency()` function for a full description. Also see the associated note for a more technical description.

```
# Test for main effects and interactions described in note.
#main_plus_interacts <- fractional_factorial_efficiency(~ maker * technical + category_rarity +
         size + age, fractional_f55_25)
```

This design supports only a single set of two-attribute interactions i.e. maker interact technical, or size interact age or age interact provenance etc. However it does not support more than two sets of two-attribute interactions: i.e. in this instance the maker interact technical and age interact provenance.

In instances where some of the stipulated effects cannot be estimated (such as above) then the D-efficiecy would be NaN and similarly the A-efficeny is zero.

Step 6: Methods to convert from factorial designs to discrete choice experiments

Modulo Method

See techincal document.

TODO: All the methods presented in step 6 should give a class of `choice_set`.

```
# dce_step <- stephens_pairing(fractional_f4521_64)
dce_modulo <- modulo_method(
  fractional_f55_25,
```

```r
  list(c(1, 1, 1, 1, 1), c(3, 3, 3, 3, 3))
)
```

Step 7: Checking for Pareto Overshadowed Cards

```r
checking_overshadow <- check_overshadow(dce_modulo)
```

Step 8: Efficacy of the Discrete Choice Design

```r
dce_modulo_efficacy <- dce_effiency(aff55, dce_modulo)
#> q is 1
#> [1] "L is 5"
#> [1] "Case 4"
#> [1] "s is 3"
#> q is 2
#> [1] "L is 5"
#> [1] "Case 4"
#> [1] "s is 3"
#> q is 3
#> [1] "L is 5"
#> [1] "Case 4"
#> [1] "s is 3"
#> q is 4
#> [1] "L is 5"
#> [1] "Case 4"
#> [1] "s is 3"
#> q is 5
#> [1] "L is 5"
#> [1] "Case 4"
#> [1] "s is 3"
#> The D-efficiency of this discrete choice experiment is 98.883 %
```

Step 9: Construct a Discrete Choice Question Frame

The function `construct_question_frame` is helpful with the final stages. It consistently converts a `choice_set` arrangement into a `data.frame`.

```r
question_table_f55 <- construct_question_frame(aff55, dce_modulo)
```

It is now time to add some useful information back to the levels. Originally these were described in Step 0, but up until this point it has been nececasury to work with only integer values. (Also just imagine if you had worked with these very long names up until this point…)

```
# levels(question_table_f4521$maker) <- c("little known", "known to specialists", "recognised",
         "famous")
# levels(question_table_f4521$technical) <- c("below average", "good", "meritorious",
         "exceptional")
# levels(question_table_f4521$category_rarity) <- c("common", "uncommon", "rare", "very rare")
# levels(question_table_f4521$size) <- c("small: under 150g", "medium: between 151g and 400g",
         "large: between 401g and 1000g", "extra large: larger than 1000g")
# levels(question_table_f4521$age) <- c("21st or 20th Century", "19th Century", "18th Century",
         "Before 18th Century")
# levels(question_table_f4521$provenance) <- c("unavailable or available, but unimportant",
         "available & important")
# View(question_table_f4521)
```

Replicating the example in Street

The purpose of this section is to replicate the … The steps follow those of the tutorial.

```
# Step 0
atttravel <- list(
  airfaire = c("0", "1"),
  travel_time = c("0", "1", "2")
)
# Step 1
travel2131 <- full_factorial(atttravel)
# Step 2
aff_travel2131 <- augment_levels(travel2131)
#> [1] "Applying B mat"
# Step 3.
# The full factorial is already so small that selecting a fraction of it would be silly.
# Therefore re-use the full factorial as the fractional fractorial.

# Step 4.
# Confirming that this is an efficient design.
fractional_travel2131 <- search_design(travel2131, travel2131)

# Step 5.
# Confirm that this design supports all interactions.
full_factorial_efficiacy <- fractional_factorial_efficiency(~ (airfaire + travel_time)^2,
         fractional_travel2131)
#> Your fractional factorial design has an A-efficiency of 100 %
#>  Your fractional factorial design has a D-efficiency of 100 %

# Step 6 & Step 7.
# Street gives two examples of choice sets.
travel_choice_set1 <- list(c("00", "11", "02"), c("10", "02", "12"))
class(travel_choice_set1) <- c(class(travel_choice_set1), "choice_set")

travel_example <- dce_effiency(aff_travel2131, travel_choice_set1)
#> q is 1
#> [1] "L is 2"
#> [1] "Case 1"
#> [1] "s is 2"
#> q is 2
#> [1] "L is 3"
#> [1] "Case 3"
#> [1] "The implied x is 1 and y is 0"
#> [1] "s is 3"
```

```
#> The D-efficiency of this discrete choice experiment is 62.996 %
# Note, if you want to rearrange the columns of the lamda matrix so that they are the same as
        Street use the following:
# lamda_street_cols  <- matrix(c(travel_example$Lamda$mat[,1],
#                                travel_example$Lamda$mat[,3],
#                                travel_example$Lamda$mat[,5],
#                                travel_example$Lamda$mat[,2],
#                                travel_example$Lamda$mat[,4],
#                                travel_example$Lamda$mat[,6]), ncol = 6)
# lamda_street_paper <- matrix(c(lamda_street_cols[1,],
#                                lamda_street_cols[3,],
#                                lamda_street_cols[5,],
#                                lamda_street_cols[2,],
#                                lamda_street_cols[4,],
#                                lamda_street_cols[6,]), ncol = 6)
# Street gives a second arrangment:
travel_choice_set2 <- list(c("00", "11", "02"), c("10", "01", "12"))
class(travel_choice_set2) <- c(class(travel_choice_set2), "choice_set")
# This version is 100% efficient.
travel_example2 <- dce_effiency(aff_travel2131, travel_choice_set2)
#> q is 1
#> [1] "L is 2"
#> [1] "Case 1"
#> [1] "s is 2"
#> q is 2
#> [1] "L is 3"
#> [1] "Case 3"
#> [1] "The implied x is 1 and y is 0"
#> [1] "s is 3"
#> The D-efficiency of this discrete choice experiment is 100 %

# Step 8
travel_questions <- construct_question_frame(aff_travel2131, travel_choice_set2,
        randomise_choice_sets = FALSE)
levels(travel_questions$airfaire) <- c("$350", "$650")
levels(travel_questions$travel_time) <- c("4 hours", "5 hours", "6 hours")
#View(travel_questions)


#Step 0
# Described in Theory
attri3261 <- list(
    starter = c("1", "2", "3"),
    main = c("1", "2", "3", "4", "5", "6"),
    dessert = c("1", "2", "3")
)
# Step 1
ff_menu <- full_factorial(attri3261)
# Step 2
aff_menu <- augment_levels(ff_menu)
#> [1] "Applying B mat"
# Step 3
nlevels <- unlist(purrr::map(ff_menu, function(x){length(levels(x))}))
#oa_feasible(36, nlevels, strength = 3)

fractional_factorial_3261_18 <- oa.design(nlevels = nlevels, columns = "min34")
# The fractional_factorial design is generated using the DoE.MIParray package.
# The following is the command to run this generation.
```

```
# The result is saved in the package.
#fractional_factorial_3261_36 <- gurobi_MIParray(36, nlevels, resolution = 3)
fractional_factorial_3261_36 <- ExpertChoice::fractional_factorial_3261_36

# Step 4
# Confirming that this is an efficient design.
colnames(fractional_factorial_3261_36) <- colnames(ff_menu)
fractional_menu_3261_36 <- search_design(ff_menu, fractional_factorial_3261_36)

# Step 5.
# Confirm that this design supports all interactions.
full_factorial_efficiacy <- fractional_factorial_efficiency(~ (.)^2, fractional_menu_3261_36)
#> Your fractional factorial design has an A-efficiency of 0 %
#>  Your fractional factorial design has a D-efficiency of 0 %
```