Consumer preferences are often measured within the framework of multiattribute decisions, for which an attribute’s importance weight should depend on the attribute’s range of levels. However, empirical studies using different ranges find that the adjustment of attribute weights occurs to a lesser degree than theoretically required. This bias is called the „range effect“. We develop a new measure, which requires fewer restrictive assumptions and is more meaningful than existing measures, to capture the strength of the range effect. We apply the proposed measure to a broad number of methods by reanalyzing the data of Sattler, Gedenk, and Hensel-Börner (2002). In addition, we measure the strength of the range effect bias for the frequently used choice-based conjoint analysis (CBC). We find that the range effect can be very powerful, particularly for self-explicated approaches, and that it is smallest for adaptive conjoint analysis (ACA) followed by CBC.

Keywords
Multiattribute decisions; preference measurement; range effects

1. Introduction

There is a long tradition of modeling multiattribute decisions and measuring decision makers’ preferences (e. g. Keeney/Raiffa 1976). Consumers’ choice of a product constitutes a key multiattribute decision in marketing. For example, consumers decide which car to buy on the basis of their evaluations of the price, brand, technical features, and color of the different cars in their choice sets. Manufacturers naturally want to know how consumers make this choice, and particularly how important products’ attributes are to decision makers, so that they can design profit-maximizing products. Conjoint analysis, self-explicated methods or the swing method are well-known examples of methods for preference measurement (e. g. Gustafsson/Herrmann/Huber 2003; von Winterfeldt/Edwards 1986). For instance, using respondents’ evaluations of stimuli, conjoint researchers estimate part-worths that reflect consumers’ preferences for attribute levels. These part-worths also contain information about attribute importance, and from them, marketers can compute importance weights.

When applying these methods to measure preferences, researchers typically assume that an alternative’s utility or value is a linear function of the conditional utilities or values of its attributes. In multiattribute decision theory, the conditional value function of each attribute is typically normalized on the interval [0,1]. The researcher then computes the overall utility of a decision alternative by weighting and adding the attribute values. Attribute weights (also called importance weights) sum to 1 across alternatives and indicate how important the attributes are for a decision maker.

When measuring these attribute weights, it is important to note that the weights should depend on the attribute range (von Nitzsch/Weber 1993). This requirement becomes intuitively evident in the context of the car example. Imagine that a consumer has to choose among three cars that differ in brand and price, with prices ranging from a minimum of €14,000 to a maximum of €16,000. The consumer has strong brand preferences, such that the importance weight for the attribute “brand” assumes a value of 0.6 versus 0.4 for price. Imagine now that the consumer has to choose between the same three brands and that prices range from €10,000 to €20,000. Given identical preferences, price should now play a greater role in the consumer’s decision relative to the brand, and the importance weight for price should be greater.

The extent to which importance weights should change for different attribute ranges is determined by the deci-
sion maker’s value function. Interestingly, several empirical studies have shown that decision makers adjust their importance weights, but not as much as they should according to their value functions. This divergence of what we empirically observe from the theoretically required adjustment of attribute weights is called the „range effect“. It implies that consumers’ importance weights cannot be measured consistently and that preference measurements may be biased. Thus, decisions based on these measurements, such as a company’s decisions on product design, are, to a certain extent, arbitrary. Note that the range effect differs from other biases that have been studied in the context of conjoint analysis, e. g. the number-of-levels effect (e. g. Verlegh/Schifferstein/ Wittink 2002) or the bias from choosing specific methods for eliciting importance weights (e. g. Jaccard/Brinberg/Ackerman 1986; Darmon/Rouziès 1991). A range effect bias does not occur because an attribute’s importance weight changes when the attribute’s range changes, but rather because the weight does not change enough.

The issue of range effects has drawn little attention in the marketing literature on preference measurement (exceptions are the papers by Verlegh/Schifferstein/Wittink 2002 and Sattler/Gedenk/Hensel-Börner 2002). Instead, most research on range effects stems from decision theory and related bodies of literature. However, range effects can have a severe impact on marketing research, including the forecasting of market shares for new products, as our study shows.

This research makes three main contributions. First, we suggest a new measure for the strength of the range effect. Compared to an existing measure (von Nitzsch/Weber 1993) our measure gives more meaningful information about the severity of the range effect and it does not require a measurement of the decision maker’s value function, which may not be error-free.

Second, we apply our new measure to a broad number of methods for eliciting attribute weights by reanalyzing the data of Sattler, Gedenk, and Hensel-Börner (2002). We find that the size of the range effect matters. In some cases the range effect bias is even stronger than the bias that stems from choosing a different method for eliciting attribute weights (e. g. adaptive conjoint analysis (ACA) versus the self-explicated method). Hence, it seems surprising that marketing researchers have worried so much about the choice of a method and so little about analyzing range effects.

Third, we for the first time investigate the range effect bias in choice-based conjoint analysis (CBC). These days, CBC is very frequently employed for measuring preferences because the realistic choice task for respondents promises high external validity compared to other methods. This holds not only for academia, but also for practice: In a survey of market research institutes in Germany, Austria, and Switzerland Sattler and Hartmann (2008) find CBC to be the most frequently used conjoint method (46 % of the applications studied). Thus, we have collected additional data, using the same decision problem and a similar sample as Sattler, Gedenk, and Hensel-Börner (2002), in order to measure the strength of the range effect for choice-based conjoint analysis (CBC).

2. Literature review

An explanation for the occurrence of range effects is provided by range theory (Vollmann 1951; Parducci 1963). It posits that judgments on attribute importance are context dependent. I. e., consumer use prices presented in the preference measurement task as reference points when making their judgments on attribute importance (e. g. Janiszewski/Lichtenstein 1999). Based on this, the literature has focused on measuring the strength of the range effect. A measure has been proposed by von Nitzsch and Weber (1993), and several studies have used it to compare different methods for preference measurement.

2.1. Von Nitzsch and Weber’s measure of range sensitivity

Von Nitzsch and Weber’s individual-level measure of the strength of the range effect is based on two sets of importance weights, which are elicited from the same respondent using the same method but with different ranges for the attribute in question. If the respondent’s value function is known, the researcher can compute how much the attribute weight should change when its range changes. The measure s compares the empirically observed adjustment of the importance weight with the theoretically expected adjustment, which is based on the value function.

More specifically, von Nitzsch and Weber (1993) assume the following additive preference model:

\[
V(a_i) = \sum_{p} g_p \cdot v_p(x_{ip})
\]  

(1)

For each attribute \( p \), there is a conditional value function \( v_p(\cdot) \), normalized on the interval [0,1]. The overall value of alternative \( a_i \) is a weighted average of the conditional values of the attributes’ outcomes, where \( x_{ip} \) is the outcome of alternative \( i \) on attribute \( p \), and the weights \( g_p \) represent each attribute’s importance. Note that in conjoint analysis, typically, the conditional value functions are not normalized, so that the estimated part-worths contain the information about attribute weights. However, this is only a difference in notation, not in substance.

The degree to which a respondent should change the importance weight of an attribute when the attribute’s range changes can be determined on the basis of the attribute’s conditional value function \( v_p \). This function is depicted in Fig. 1.

Assume that the original range of attribute \( p \) is \([x_{ip}^-, x_{ip}^+]\) or \( R_p \). Because the conditional value function \( v_p \) is nor-
Range sensitivity $s$ thus is computed by comparing $m$ with the change in the importance weight that has been measured empirically, $m_{\text{emp}}$:

$$s = \frac{m_{\text{emp}} - 1}{m - 1}$$

Figure 1: Theoretically required change in importance weight

The measure $s$ has laid the groundwork for several interesting studies on range effects and thus contributed much to the literature. Nevertheless, it suffers from two key drawbacks. First, it is not in itself very meaningful to managers. Although it can be interpreted as the percentage of the change in range that a decision maker has taken into account (von Nitzsch and Weber 1993), it is hard to tell what its consequences will be in a practical business situation. For example, should managers be worried about $s = 0.4$? Managers are not used to thinking in adjustments of importance weights. Also, whether a particular $s$ value is cause for concern depends on how important the attribute in question is for a consumer’s decision, and $s$ does not capture that. Second, $s$ is based on a respondent’s individual conditional value function $v_{\text{p}}$. The researcher needs to know the value function to compute how much the importance weights should change. In our example, we have simply assumed that the conditional value function is linear. However, real value functions are typically non-linear, and need to be measured. Thus, in order to compute $s$, a researcher needs to elicit both the importance weights for different ranges, and the value function for the attribute whose range varies. Given that research on range effects focuses on biases in preference measurement, it is rather bold to assume that the conditional value function can be measured without bias. We therefore suggest a new measure of range sensitivity that does not require measurement of the value function and is more meaningful to managers.

2.2. Strength of the range effect for different methods of preference measurement

Results of previous studies on range effects are summarized in Tab. 1. The Table includes values of $s$ whenever it has been computed in these studies, and indicates whether $s$ is larger than 0 or smaller than 1 for studies that do not explicitly use the $s$ measure.

We can learn three things from Tab. 1. First, respondents adjust importance weights when an attribute’s range changes. All values for $s$ are greater than 0 with one exception. Second, respondents do not adjust importance weights as much as they should; the values for $s$ reported in the studies — again with one exception — are clearly smaller than 1. Third, the self-explicated (or direct ratio) method seems to perform worse than the swing or conjoint approaches, in particular ACA. Values between 0.07 and 0.28 for $s$ with the self-explicated method indicate only a very small adjustment of importance weights. Note that the range effect has not been studied, yet, for choice-based conjoint analysis (CBC) which is very frequently used in academia as well as in practice (Sattler/Hartmann 2008) to measure preferences.

3. A new measure for range sensitivity

We call our measure „choice share deviation“ (CSD) because it is based on how much choice shares change when an attribute’s range changes. Similar to $s$, our CSD measure is based on two sets of importance weights elicited using the same method for different ranges of an attribute. Also, like von Nitzsch and Weber (1993), we cannot say which of the two measurements is correct but can only measure differences in importance weights. The stronger these differences, the stronger the range effect is. Our measure differs from $s$ in that it is not computed at the individual level but rather across respondents. As a consequence, CSD does not allow us to identify consumers who are plagued more or less by the range effect, but the measure becomes more robust. The procedure for computing CSD is as follows:

- We forecast each respondent’s choice from a choice set that contains several alternatives, such as various products described by their attributes. We assume that a respondent chooses the alternative that has the highest value for him (first choice). We make two forecasts for each respondent on the basis of the two different sets of importance weights elicited with different ranges of one of the attributes.
- For each range (large and small), we aggregate across respondents by computing choice shares, i.e., the per-
percentage of respondents that chooses each alternative. For each alternative in the choice set, we then compute the difference between the choice shares for both ranges. We finally compute the mean absolute deviation (MAD) between the choice shares for the two ranges across all alternatives.

- Because the deviation of the choice shares may depend on the alternatives in the choice set, we draw several sets of choice alternatives from all possible alternatives and compute the MAD for each. Our CSD measure is the mean of the MAD values across choice sets.

Although CSD builds on the assumption that consumers have a conditional value function for the attribute in question, we do not measure that value function, nor do we use it to predict how attribute weights should be adjusted. Rather, we capture the theoretically required adjustment of attribute weights through predictions of choice shares. If respondents adjust their weights correctly, choice shares should be the same for both attribute ranges, irrespective of the shape of the value function. The larger the differences in choice shares, the stronger the range effect is. Furthermore, CSD is particularly meaningful to managers because choice shares have a strong relationship to market shares. If awareness and distribution are the same for all brands, choice shares are very similar to market shares. In that case, CSD tells managers by how many percentage points the predicted market shares differ on average when different ranges are used to elicit attribute importance [1]. Also, CSD captures how well consumers adjust their importance weights as well as the absolute importance of the attribute in question. CSD will only be large for an important attribute.

### 4. Methods for measuring importance weights

In our study we reanalyse the data set of Sattler, Gedenk, and Hensel-Börner (2002), which covers nine methods for preference measurement. In addition, we have collected new data for choice-based conjoint analysis (CBC). In total, we study 10 methods for eliciting importance weights that we now describe in more detail. We use compositional and decompositional methods for measuring preferences as well as hybrid approaches, namely adaptive conjoint analysis (ACA, Johnson 1987) and computerized customized conjoint analysis (CCC, Hensel-Börner/Sattler 1999; Srinivasan/Park 1997). In the study by Sattler and Hartmann (2008) ACA is second among conjoint methods in frequency of commercial use (34 % of applications). ACA and CCC measure preferences in three stages. The first stage consists of a self-explicated approach, in which ACA uses absolute ratings of attribute importance, and CCC asks for ratings relative to the most important attribute. Specifi-

<table>
<thead>
<tr>
<th>Study</th>
<th>Decision Problem</th>
<th>Method</th>
<th>Conjoint</th>
<th>Swing</th>
<th>Self-Explicated</th>
<th>Rating</th>
<th>Ranking</th>
<th>Pairs</th>
<th>Trade-Off</th>
<th>ACA</th>
<th>CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creyer and Ross, 1988</td>
<td>purchase of a car</td>
<td>Conjoint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goldstein, 1990</td>
<td>choice of an apartment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beattie and Baron, 1991</td>
<td>choice of basketball players, employees</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>von Nitzsch and Weber, 1993</td>
<td>choice of a job</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fischer, 1995</td>
<td>choice of a job</td>
<td>ACA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pöyhönen and Hämäläinen, 1998</td>
<td>choice of a job</td>
<td>CCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verlegh et al., 2002</td>
<td>choice of a TV set</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sattler et al., 2002</td>
<td>choice of a bus trip</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$s =$ average range sensitivity
* significant (p < 0.5)
* range sensitivity is greater for conjoint analysis than for the self-explicated method

| Table 1: Range sensitivity in previous empirical studies |
cally, ACA points out the difference between the worst and the best outcome of an attribute and asks „If two bus trips were the same in all other ways, how important would this difference be to you?“ For CCC a value of 10 is assigned to the most important attribute. Then respondents are asked to evaluate the other attributes relative to the most important one and assign importance ratings between 0 and 10. We can compute a set of importance weights for each of these self-explicated approaches, which we refer to as SE-ACA and SE-CCC. Because SE-ACA puts more emphasis on the range of the attribute than SE-CCC, we expect SE-ACA to suffer a little less from the range effect. Secondly, respondents confront a conjoint task. With ACA, the task consists of 13 paired comparisons of products described by two to three attributes. For CCC, respondents have to rank nine products described by all four attributes. The second stage also results in a set of importance weights for each method, which we call PAIR-ACA and RANK-CCC. Finally, the third stage is the calibration stage, in which we combine the results from the previous two stages in a weighted average. The weights are determined on the basis of the respondents’ answers to a third preference measurement task. With ACA, respondents indicate their purchase probabilities for four products described by all four attributes, whereas with CCC, they rank order another set of nine products described by all four attributes. We refer to the importance weights, which are the end result of the two hybrid methods, as ACA and CCC. In summary, each of the hybrid methods generates three sets of importance weights, one for each stage, which results in six sets of importance weights (i.e., SE-ACA, SE-CCC, PAIR-ACA, RANK-CCC, ACA, and CCC).

We also applied four paper-and-pencil methods. The first is a traditional self-explicated approach, called SE-TRAD, for which we asked the same questions as in the CCC, but used a paper-and-pencil response format instead of a computer-assisted interview. Thus, we can check whether computer assistance has an impact on the extent of the range effect. Differences could occur because some respondents may find a computerized interview more stimulating, and thus may pay closer attention because some respondents may find a computerized interview more stimulating, and thus may pay closer attention. The second paper-and-pencil approach (SE-NOTE) uses the same questions but adds a notice that tries to alert respondents to the range effect by urging them to take ranges into account when determining attribute weights. The notice reads as follows:

Please note that the importance of an attribute depends on how much the best and the worst outcome of this attribute differ. For example, the price of a bus trip should be less important compared to other attributes of the trip, if prices of the trips offered differ by DM 50, compared to situations where the difference between the most and the least expensive trip is DM 500.

This modification of the self-explicated approach was suggested by Sattler, Gedenk, and Hensel-Börner (2002), to check if a simple notice can alleviate the problem of the range effect bias for self-explicated methods. The studies in Tab. 1 show that the range effect is particularly strong for this method. At the same time, many studies find that the self-explicated approach is often as good as conjoint methods with respect to predictive validity (e.g., Sattler/Hensel-Börner 2003, Srinivasan/Park 1997). Thus, if an easy solution could be found for the range effect problem for the self-explicated approach, researchers could use this method instead of the more complicated conjoint analysis or hybrid approaches.

Third, we used the swing method (SWING) (von Winterfeldt/Edwards 1986). This method requires respondents to indicate the best and worst outcomes of each attribute. Based on the alternative that presents all attributes at their worst outcome, respondents rate the improvements, one at a time, achieved when attributes are changed to their best outcome.

Finally, we applied choice-based conjoint analysis (CBC). We used the Sawtooth software to generate 12 choice sets, each of which includes three products, and estimated individual-level parameters with Hierarchical Bayes.

5. Data

5.1. Decision problem

Our study uses survey data collected from consumers in Germany (Sattler/Gedenk/Hensel-Börner 2002). The choice of a four-day vacation trip by bus to Paris was used as the multiattribute decision. The choice problem is relevant to many German consumers, and worked well in another conjoint study (Baier/Säuberlich 1997). We used four attributes that Baier and Säuberlich found to be important, namely hotel category (with the attribute levels 1, 2, or 3 stars), visit to Euro-Disney (yes or no), sightseeing program (no program, tour of the city, or tour of the city and cultural events), and price. For the attribute price (measured in DM = Deutsche Mark), we used two ranges: a small one, with prices ranging from DM 320 to DM 480 (and an intermediate level of DM 400), and a large one, with prices ranging from DM 250 to DM 550 (and also an intermediate level of DM 400). Sattler, Gedenk, and Hensel-Börner (2002) found these ranges to be realistic in exploratory interviews with consumers and tour operators.

5.2. Data collection

In this study we combine two survey data sets. First, we use the data set collected by Sattler, Gedenk, and Hensel-Börner (2002) through personal interviews in a mid-sized German town, which covers nine different methods for eliciting importance weights. Respondents were chosen on the basis of quota sampling: 50 % men and 50 % older than 35 years of age. Second, we have collected additional data on CBC. Since Sattler, Gedenk,
and Hensel-Börner find hardly any differences in range sensitivity between students and non-students, we have surveyed students at the University of Hamburg for the second data set [2]. Both surveys used the same decision problem, as described in the previous sub-section. Note that the data for CBC was collected a few years later than the first data set. Thus, prices were translated into Euro (€). We believe that the two the data sets are comparable because prices for bus trips have not changed much. Also, we used the traditional self-explicated approach in both surveys and find extremely similar results, confirming stability over time.

In both surveys, the attribute range varied within subjects, whereas the methods varied between subjects. In total, we have collected data from six subsamples regarding the following methods: ACA (resulting in SE-ACA, PAIR-ACA, and ACA), CCC (resulting in SE-CCC, RANK-CCC, and CCC), SE-TRAD, SE-NOTE, SWING, and CBC.

The first task of each respondent was to answer questions that pertained to measuring attribute importance on the basis of either the small or the large price range. This was followed by questions on respondents’ demographic characteristics. Next, the second preference measurement task applied the identical method for the other price range. The order in which the small and the large range were presented varied randomly. We also measured each respondent’s conditional value function for the price attribute to compute range sensitivity s, as suggested by von Nitzsch and Weber (1993). For this purpose, a price range of DM 200–600 was used, which covers the DM 250–500 range used to elicit preferences. Respondents were told that the value of the price „DM 200“ equaled 100, and were then asked to assign values to prices of DM 300, 400, 500, and 600. To compute the values for prices between these data points, we used linear interpolation.

Each of the six subsamples consists of approximately 80 respondents, which results in a total of 497 respondents and 813 observations (because ACA and CCC generate three measures each). We only included respondents who had at least some interest in a bus trip to Paris.

Following Sattler, Gedenk, and Hensel-Börner (2002), we check the face validity of our results by computing von Nitzsch and Weber’s (1993) range sensitivity index s for each respondent. The distribution for s includes some values substantially smaller than 0 and larger than 1. These implausible values occur for all methods, and they indicate that some respondents did not understand the task or did not pay sufficient attention to completing it. Overall, our sample contains 9.7 % of observations with s values below –1 and above 3 (theoretical value of 1 ± 2). We treat them as outliers and eliminate them. The data set thus contains 734 observations. Note, that there is no clear limit beyond which an s value is implausible, so we test for the robustness of our results by using an alternative limit. When we exclude values of s between –0.25 and 2.25 (theoretical value of 1 ± 1.25), we find that the values of s for the ten methods are very similar and evince an across-sample correlation of 0.99. This indicates that our results are robust to outliers.

5.3. Measurements of range sensitivity

We compute range sensitivity s as suggested by von Nitzsch and Weber (1993). To compute our choice share deviation (CSD) measure of range sensitivity, we simulate choice shares for both the small and large ranges for a set of three products. The small and the large price range contain three price levels each, with the medium level being the same. We have derived part-worths for these price levels with our 10 different methods for preference measurement, and we use interpolation and extrapolation to compute part-worths for the remaining two price levels for both the small and the large range [3]. The three products we use for the simulation are randomly drawn from 90 possible products (2 Euro-Disney × 3 hotel × 3 sight-seeing × 5 price outcomes). We draw again whenever one of the three products is dominated by another (with the assumption that lower prices and more stars for hotels would be considered better). We repeat the analysis for 100 sets of three products and compute CSD as the mean of the 100 MAD values. Thus, CSD indicates the mean difference in predicted choice shares when predictions are based on different attribute ranges for price. It is a meaningful measure that does not require a measurement of the value function.

We compute our CSD measure based on a within-subjects experimental design. Unlike range sensitivity s, however, our measure could also be derived from between-subjects data, if a researcher were worried about conditioning effects. In our sample we find that CSD generates similar results with data from within- and between-subjects designs. To demonstrate this, we compute CSD values using only the first preference measurement from each respondent (randomly based on either the large or small range). The results are similar to those based on the within-subjects design, and the CSD measures for both approaches evince a correlation of 0.77. We devote the remainder of this article to presenting the results from our within-subjects design to achieve comparability with s and higher reliability due to the larger data set.

6. Results

In Tab. 2, we show the strength of the range effect for all ten investigated methods, ordered according to CSD. The methods with the smallest CSD value, and thus the least severe range effect, appear at the top of the table.

Tab. 2 displays importance weights for the attribute price given the small and large ranges, computed for each respondent and then averaged across respondents (see also Sattler/Gedenk/Hensel-Börner 2002). The importance weight of price is greater when the large price
Table 2: Range effects for different methods

<table>
<thead>
<tr>
<th>Method</th>
<th>No. of Observations</th>
<th>Importance Weight for Price&lt;sup&gt;a&lt;/sup&gt;</th>
<th>s&lt;sup&gt;a&lt;/sup&gt;</th>
<th>CSD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Small Range</td>
<td>Large Range</td>
<td></td>
</tr>
<tr>
<td>ACA</td>
<td>73</td>
<td>0.26 (0.13)</td>
<td>0.33 (0.13)</td>
<td>0.73 (0.84)</td>
</tr>
<tr>
<td>CBC</td>
<td>77</td>
<td>0.30 (0.18)</td>
<td>0.36 (0.21)</td>
<td>0.56 (0.79)</td>
</tr>
<tr>
<td>RANK-CCC</td>
<td>60</td>
<td>0.31 (0.23)</td>
<td>0.36 (0.23)</td>
<td>0.47 (0.76)</td>
</tr>
<tr>
<td>SWING</td>
<td>67</td>
<td>0.20 (0.10)</td>
<td>0.23 (0.11)</td>
<td>0.45 (0.76)</td>
</tr>
<tr>
<td>PAIR-ACA</td>
<td>64</td>
<td>0.26 (0.11)</td>
<td>0.33 (0.12)</td>
<td>0.70 (0.85)</td>
</tr>
<tr>
<td>SE-NOTE</td>
<td>88</td>
<td>0.27 (0.10)</td>
<td>0.30 (0.10)</td>
<td>0.28 (0.53)</td>
</tr>
<tr>
<td>SE-ACA</td>
<td>72</td>
<td>0.28 (0.09)</td>
<td>0.31 (0.09)</td>
<td>0.26 (0.54)</td>
</tr>
<tr>
<td>CCC</td>
<td>74</td>
<td>0.31 (0.12)</td>
<td>0.32 (0.11)</td>
<td>0.15 (0.44)</td>
</tr>
<tr>
<td>SE-CCC</td>
<td>78</td>
<td>0.30 (0.10)</td>
<td>0.31 (0.10)</td>
<td>0.09 (0.36)</td>
</tr>
<tr>
<td>SE-TRAD</td>
<td>81</td>
<td>0.29 (0.09)</td>
<td>0.30 (0.09)</td>
<td>0.07 (0.29)</td>
</tr>
</tbody>
</table>

<sup>a</sup> mean (standard deviation)

The range is used to measure attribute importance in all methods. This is good news because it indicates that respondents change their attribute weight when they view a different attribute range in the preference measurement task. The bad news, however, is that the change in attribute weight is rather small for some methods. The same good and bad news is reflected by von Nitzsch and Weber’s (1993) measure of range sensitivity s, for which larger values indicate a less severe range effect (see also Sattler/Gedenk/Hensel-Börner 2002). The average s is greater than 0 but smaller than 1 for all methods. This implies that respondents adjust attribute weights – but not as much as they theoretically should. Similar to previous studies, we find that s is particularly small for self-explicated methods and greater for swing and traditional conjoint methods, which indicates that the latter are less plagued by the range effect than the former.

As the primary focus of this study, Tab. 2 contains our CSD measure of range sensitivity, for which larger values indicate a stronger range effect. Comparing s and CSD, we find a very similar ordering of methods. The only substantive difference appears with respect to the paired comparison task within ACA (PAIR-ACA). This method performs rather well with s and achieves the second-best value of 0.70. However, it performs less well with CSD, achieving only the fifth-best CSD value of 7.41. This inconsistency could be due to several reasons: (1) s is computed at the individual level, whereas CSD is an aggregate measure; (2) s uses a measurement of the value function, whereas CSD does not; and (3) s uses only the importance weight of price, whereas CSD uses the preference information for all attributes. Nevertheless, the overall similarity between the two measures is very strong, and the correlation between them across the nine methods is -0.84, which suggests high convergent validity.

Hereafter, we rely on CSD to compare the ten methods of preference measurement. CSD indicates the average percentage point change in the choice share prediction when a different price range is used. For example, when measuring preferences with ACA, we find an average difference of 3.69 percentage points in choice shares measured with the small versus the large range. The difference is 10.14 percentage points for the traditional paper-and-pencil self-explicated approach (SE-TRAD). This represents a substantial deviation relative to an average market share of 33% (given that there are always three products in the market in our simulation). That is, the range effect matters.

To demonstrate the importance of the range effect further, we also compute choice share deviations that result from applying different methods. The CSD measures in Tab. 2 compare choice shares for different ranges within method. We also compare choice shares for different methods within range. For this exercise, we choose the best and worst methods from Tab. 2, namely, ACA and the paper-and-pencil self-explicated approach (SE-TRAD). We find a difference of 8.56 percentage points in choice share prediction between ACA and SE-TRAD for the small range and 8.97 points for the large range. The bias from the range effect with a self-explicated approach is therefore greater than the bias that results from applying a different method. In turn, it seems surprising that marketing researchers have been worrying a lot about their choice of a method, but less about the choice of attribute range, which is just as important or, for some methods, even more so.
For further comparison of the ten methods, we must consider whether the differences we observe in Tab. 2 are significant. In Tab. 3, we therefore indicate the significance of pairwise differences, which result from Wilcoxon tests that compare MAD values for the 100 product sets between two methods at a time. The CSD values of all methods are repeated at the top of the table to facilitate interpretation.

Tab. 3 reveals several interesting insights. First, the range effect is particularly strong for self-explicated methods. The four self-explicated approaches are among the five methods with the greatest CSD values. This confirms findings from previous studies and thus demonstrates high face validity of our results. Second, there is no significant difference between SE-TRAD and SE-CCC, which indicates that it does not matter whether the self-explicated approach uses paper-and-pencil or a computerized version. Using computer support thus does not stimulate respondents enough to alleviate the range effect bias. Third, it helps to alert respondents to the range effect. SE-ACA suffers significantly less from the range effect than the two traditional approaches SE-TRAD and SE-CCC. This is probably due to the fact that the state-of-the-art self-explicated approach in ACA asks the attribute importance question as the improvement from the least to the most preferred level of the attribute. It helps even more to add a simple notice to the self-explicated approach, telling respondents about the range effect. We find that CSD decreases from 10.14 for SE-TRAD to 7.84 for SE-NOTE, the identical method with the notice. SE-NOTE is significantly and substantially better than both SE-TRAD and SE-ACA. Unfortunately, the self-explicated approach with the notice still suffers more strongly from the range effect bias than more complex methods for eliciting importance weights. That is, the notice helps, but it does not solve the problem. We therefore suggest that researchers who use the self-explicated approach add a notice to alert the respondents to the range effect; but we also remind researchers that the range effect bias will be even smaller if they use other methods to measure preferences.

Fourth, SWING performs significantly better than all the self-explicated approaches investigated. It is not significantly better or worse than RANK-CCC and CBC, and is outperformed only by ACA. Thus, for researchers looking for a simple method, SWING may be a good choice.

Fifth, among the hybrid approaches ACA outperforms CCC. Results are not quite clear when we look only at the conjoint part of these methods. Traditional ranking (RANK-CCC) has a significantly better CSD value than paired comparisons (PAIR-ACA), but range sensitivity is better for PAIR-ACA than for RANK-CCC. Overall it is clear, however, that ACA performs substantially better than CCC. The calibration in the third stage of ACA seems to work extremely well, even though both parts of ACA have rather large CSD values when used by themselves. This results in the best CSD value for overall ACA importance weights.

Sixth, ACA also significantly outperforms choice-based conjoint analysis (CBC). While the difference in CSD is not huge, it is substantive: The mean absolute error in predicted market share is 3.7 percentage points for ACA compared to 4.6 percentage points for CBC. I. e., although CBC uses choice questions to elicit importance weights and is thus particularly realistic, it does not

<table>
<thead>
<tr>
<th></th>
<th>ACA</th>
<th>CBC</th>
<th>RANK-CCC</th>
<th>SWING</th>
<th>PAIR-ACA</th>
<th>SE-NOTE</th>
<th>SE-ACA</th>
<th>CCC</th>
<th>SE-CCC</th>
<th>SE-TRAD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.69</td>
<td>4.60</td>
<td>4.79</td>
<td>5.01</td>
<td>7.41</td>
<td>7.84</td>
<td>9.13</td>
<td>9.31</td>
<td>10.07</td>
<td>10.14</td>
</tr>
<tr>
<td>ACA</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>CBC</td>
<td>n.s.</td>
<td>n.s.</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>RANK-CCC</td>
<td>n.s.</td>
<td>n.s.</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>SWING</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>PAIR-ACA</td>
<td>n.s.</td>
<td>n.s.</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>SE-NOTE</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>SE-ACA</td>
<td>n.s.</td>
<td>n.s.</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>CCC</td>
<td>***</td>
<td>n.s.</td>
<td>**</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>SE-CCC</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Wilcoxon test: n.s. = not significant
* p < 0.1,  ** p < 0.05,  *** p < 0.01

Table 3: Significance of differences in CSD between methods
avoid the range effect quite as well as ACA. This is an interesting result against the background that choice-based conjoint procedures have gained considerable attention in the academic literature in the last decade.

7. Summary and conclusions

We have developed a new measure of the strength of the range effect. Our measure has two key advantages compared with range sensitivity s, as suggested by von Nitsch and Weber (1993): It does not require the measurement of the conditional value function, and it is more meaningful. In addition, it does not require within-subjects experimental data, but can also be computed based on between-subjects data.

We have applied our measure to ten methods for eliciting attribute importance. Our results mostly confirm those of Sattler, Gedenk, and Hensel-Börner (2002). The two measures for range sensitivity yield contradictory results for the comparison of conjoint analysis with ranking versus paired comparisons. However, they consistently show the superiority of ACA as well the inferiority of simple self-explicated approaches with respect to the range effect. In addition to Sattler, Gedenk, and Hensel-Börner (2002), our new measure now allows us to show that the range effect bias is indeed important and that marketing researchers should pay close attention to it. Also, we have generated new insights about the strength of the most frequently used conjoint method CBC. In summary, we have found the following:

- Attribute range has a strong impact on predicted choices.
- The range effect is strongest for self-explicated methods.
- Adding a notice that alerts respondents to the problem alleviates it to some degree.
- It does not make a difference whether the task is presented using paper and pencil or on a computer.
- SWING suffers less from the range effect bias than do self-explicated approaches and is significantly outperformed only by ACA.
- ACA is superior to CCC among hybrid methods.
- ACA is also superior to CBC. Overall, ACA suffers least from the range effect.

We can draw several conclusions from these results. We find that the range effect matters! While previous research measures the relative strength of the effect compared across methods, our new measure CSD has enabled us to also measure its absolute importance. We find that the difference in predicted choice shares from using a different range can be larger than the difference from applying a different method. Thus, the range effect bias must be considered by marketing researchers, who so far have paid almost no attention to this problem. Our results highlight the importance of choosing a realistic attribute range when measuring attribute importance.

Further, we have generated insights with regard to the choice of methods for measuring attribute weights. We find that choice of method clearly matters with respect to range effects. In particular, we would like to highlight three findings. First, our study confirms previous results that simple self-explicated approaches suffer strongly from the range effect. However, Sattler, Gedenk, and Hensel-Börner (2002) suggested a modification of the self-explicated approach, and we have demonstrated its usefulness. Market researchers who want to use self-explicated approaches may find it helpful to add a notice in their survey that alerts respondents to the problem of the range effect. Thus, researchers can mitigate range effects by carefully instructing respondents. Second, among the simple approaches, we find that swing outperforms self-explicated methods with respect to the range effect. This method has not received much attention in the marketing literature. However, our research suggests it may be an interesting alternative to self-explicated methods. Third, ACA is the best among many methods with respect to the range effect, outperforming even CBC. Both ACA and CBC are used by a high number of researchers and managers. By studying the range effect for these methods, we have uncovered an advantage of ACA that could help guide marketing researchers in their choice of a method.

It would be interesting for further research to investigate whether a notice that alerts respondents to the range effect can also improve other methods besides the self-explicated approach. If this is the case, it might make ACA even more robust with respect to the range bias. Also, more research into causes of the range effect is needed to help design better methods for measuring preferences. Meanwhile, our results show which methods suffer least from range effect bias and consequently assist marketing researchers in choosing an appropriate method.

Notes

[1] Note that – unlike s – CSD does not indicate whether consumers adjust their importance weights too much or too little. However, we know from previous research that on average the adjustment is smaller than theoretically required. Also, we do not consider this information crucial, given that we do not know which of the two ranges is „correct“. [2] Sattler, Gedenk, and Hensel-Börner (2002) find differences between students and non-students only for ranking-based conjoint analysis, which requires respondents to perform a very complicated task. This is not the case for CBC. Thus, we do not expect any biases from sampling only students for CBC. [3] The small range includes the price levels 320, 400, and 480 DM. We use extrapolation to generate utilities for 250 and 550 DM (based on the difference in utility between 320 and 480 DM). The large range includes the price levels 250, 400, and 550 DM. We use interpolation to generate utilities for 250 and 480 DM (based on the differences between 250 and 400 and between 400 and 550 DM).
References


