This item is the archived peer-reviewed author-version of:

Homogeneous versus heterogeneous designs for stated choice experiments : ain't homogeneous designs all bad?

Reference:

Kessels Roselinde.- Homogeneous versus heterogeneous designs for stated choice experiments : ain't homogeneous designs all bad?
Journal of Choice Modelling - ISSN 1755-5345 - 21(2016), p. 2-9
Full text (Publisher's DOI): <https://doi.org/10.1016/J.JOCM.2016.08.001>
To cite this reference: <http://hdl.handle.net/10067/1387130151162165141>

Homogeneous versus heterogeneous designs for stated choice experiments: Ain't homogeneous designs all bad?

Roselinde Kessels

Universiteit Antwerpen
Department of Economics
& StatUa Center for Statistics
Prinsstraat 13
2000 Antwerpen, Belgium
`roselinde.kessels@uantwerpen.be`

Abstract

Some recent attempts on constructing heterogeneous designs for stated choice experiments where different respondents or groups of respondents get different subdesigns have proven successful. Compared to homogeneous designs where all respondents get the same choice sets, heterogeneous designs allow for more variation in the attribute levels resulting in a larger amount of information on the respondents' preferences. Homogeneous designs have remained popular, however, because they are easier to generate and implement. In this paper, the question is raised about when homogeneous designs perform almost as well as heterogeneous designs under the Bayesian multinomial logit design framework. A simulation study is presented to identify the situations where the losses in estimation efficiency from using a homogeneous design are small and where they are large. When the residual degrees of freedom from using a homogeneous design are large and, to a lesser extent, the number of attributes and attribute levels are small, the efficiency losses are negligible and the use of a homogeneous design can be justified.

Keywords: Stated choice experiments, Homogeneous design, Heterogeneous design, Different subdesigns, Bayesian D -optimality, D_B -efficiency

1 Introduction

In recent years, some research on discrete or stated choice (SC) experimental design has focused on the efficient construction of heterogeneous designs where different respondents or groups of respondents get different designs, i.e. subdesigns from a larger design. Sándor and Wedel (2005) were the first to propose the use of heterogeneous designs for SC experiments. Even though the number of subdesigns was small in the heterogeneous designs they constructed, Sándor and Wedel (2005) showed that heterogeneous designs provide more information on the respondents' preferences compared to homogeneous designs where each respondent gets the same choice sets. They presented evidence by comparing homogeneous versus heterogeneous multinomial and mixed logit designs in a locally or Bayesian optimal design framework. The construction of heterogeneous locally optimal mixed logit designs is computationally intensive. Liu and Tang (2015) outperformed the approach of Sándor and Wedel (2005) for generating these designs by a new approach that makes it practical to generate a completely heterogeneous design where each respondent is given a unique subdesign.

Various SC applications using heterogeneous Bayesian designs have been carried out by Kessels et al. (2015a) who studied preferences for software output displays, by Kessels et al. (2015b) and Luyten et al. (2015) who studied health care related preferences and by Verhetsel et al. (2015, 2016) and Kupfer et al. (2016) who studied preferences in transport and housing. In each of these applications, the heterogeneous designs consist of two to eight subdesigns or surveys that were evenly spread over all respondents. The allocation of the choice sets to surveys, referred to as blocking, was quasi-random in the sense that each survey had to satisfy a similar structure of attributes whose levels were constant. The choice sets in question consisted of so-called partial profiles in order to cope with larger numbers of attributes. In the case of full profile designs, where the levels of all attributes are allowed to vary, blocking can be done fully randomly (Reed Johnson et al. 2013). Besides (quasi-)random blocking, blocking has also been done by minimizing the average correlation between an additional blocking variable and all other design attributes. This ensures that attribute level balance is maintained within each survey as much as possible. Bliemer and Rose (2011) applied such strategy for the creation of heterogeneous Bayesian designs using the choice design software Ngene (ChoiceMetrics 2014).

Despite the dissemination of heterogeneous designs for SC experiments, homogeneous designs have still been frequently used because they are easier to generate and implement (see, e.g., Sándor and Wedel (2001, 2002), Rose et al. (2008), Hess and Rose (2009), Yu et al. (2009), Bliemer and Rose (2010), among others). Therefore, the question is raised to what extent homogeneous designs are underperforming compared to heterogeneous designs. Ain't homogeneous designs all bad? To answer this question, the estimation efficiency of a homogeneous design versus a heterogeneous design is derived and computed for various scenarios presented in a simulation study. Underlying the design generation is the Bayesian multinomial logit (MNL) design framework which provides the building blocks for many other discrete choice designs. It is described in Section 2 and the simulation study is presented in Section 3. Finally, Section 4 discusses the results and concludes.

2 The Bayesian MNL design framework

The MNL model assumes that respondents to a SC experiment belong to a target group of decision makers with homogeneous preferences. The model employs random utility theory which describes the utility that a respondent attaches to profile j ($j = 1, \dots, J$) in choice set s ($s = 1, \dots, S$) as the sum of a systematic and a stochastic component:

$$U_{js} = \mathbf{x}'_{js}\boldsymbol{\beta} + \varepsilon_{js}. \quad (1)$$

In the systematic component $\mathbf{x}'_{js}\boldsymbol{\beta}$, $\mathbf{x}_{js}$ is a $k \times 1$ vector containing the attribute levels of profile j in choice set s . The vector $\boldsymbol{\beta}$ is a $k \times 1$ vector of parameter values representing the effects of the attribute levels on the utility. This parameter vector is the same for every respondent. The stochastic component ε_{js} is the error term, which is assumed independently and identically extreme value distributed. Therefore, the MNL probability that a respondent chooses profile j in choice set s is the closed-form expression

$$p_{js} = \frac{\exp(\mathbf{x}'_{js}\boldsymbol{\beta})}{\sum_{t=1}^J \exp(\mathbf{x}'_{ts}\boldsymbol{\beta})}, \quad (2)$$

where $\boldsymbol{\beta}$ can be estimated using a maximum likelihood approach.

The construction of an optimal design $\mathbf{X} = [\mathbf{x}'_{js}]_{j=1,\dots,J;s=1,\dots,S}$ for estimating $\boldsymbol{\beta}$ in the MNL model (2) requires some form of maximization of the Fisher information matrix with respect to $\mathbf{X}$. The information matrix is defined as

$$\mathbf{M}(\mathbf{X}, \boldsymbol{\beta}) = \sum_{s=1}^S \mathbf{X}'_s (\mathbf{P}_s - \mathbf{p}_s \mathbf{p}'_s) \mathbf{X}_s, \quad (3)$$

with $\mathbf{X}_s = [\mathbf{x}'_{js}]_{j=1,\dots,J}$ the submatrix of $\mathbf{X}$ corresponding to choice set s , $\mathbf{p}_s = [p_{1s}, \dots, p_{Js}]'$ and $\mathbf{P}_s = \text{diag}[p_{1s}, \dots, p_{Js}]$. To maximize the information content in (3), a robust design strategy is to maximize the Bayesian $\mathcal{D}$ -optimality criterion or $\mathcal{D}_B$ -optimality criterion, which is defined as

$$\mathcal{D}_B = \int_{\mathcal{R}^k} \log |\mathbf{M}(\mathbf{X}, \boldsymbol{\beta})| \pi(\boldsymbol{\beta}) d\boldsymbol{\beta}. \quad (4)$$

In this expression, $\pi(\boldsymbol{\beta})$ stands for a prior distribution of likely parameter values used to compute the probabilities p_{js} for obtaining the information matrix (3). The $\mathcal{D}$ -optimality criterion, i.e. the logarithmic transformation of the determinant of the information matrix, is averaged over the prior distribution. This approach is known as the Bayesian $\mathcal{D}$ -optimal design approach (Sándor and Wedel 2001). Often, the prior distribution is the multivariate normal distribution, $\mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$, with prior mean $\boldsymbol{\beta}_0$ and prior variance-covariance matrix $\boldsymbol{\Sigma}_0$. The $\mathcal{D}$ -optimality criterion can be integrated over this distribution using the fast and accurate quadrature scheme by Gotwalt et al. (2009) and Gotwalt (2010) which requires only $O(k^2)$ function evaluations. The design that maximizes the $\mathcal{D}_B$ -criterion is the $\mathcal{D}_B$ -optimal design. Such design has been shown to perform well for a broad range of true parameter vectors $\boldsymbol{\beta}$ (Kessels et al. 2011b,c; Rose 2011).

In practice, the $\mathcal{D}_B$ -criterion value is approximated by drawing d prior parameter values $\boldsymbol{\beta}^i$, $i = 1, \dots, d$, from $\pi(\boldsymbol{\beta})$, and computing

$$\tilde{\mathcal{D}}_B(\mathbf{X}) = \frac{1}{d} \sum_{i=1}^d \log |\mathbf{M}(\mathbf{X}, \boldsymbol{\beta}^i)|. \quad (5)$$

A design $\mathbf{X}$ is homogeneous insofar as it is assigned to a group of T respondents. Let $\mathbf{X}_{\mathbf{T}} = [\mathbf{X}_{(t)}]_{t=1, \dots, T}$ be the stacked design matrix with T replications of $\mathbf{X}$. The $\tilde{\mathcal{D}}_B$ -criterion value (5) of $\mathbf{X}_{\mathbf{T}}$ can be defined as

$$\tilde{\mathcal{D}}_B(\mathbf{X}_{\mathbf{T}}) = \frac{1}{d} \sum_{i=1}^d \log |\mathbf{M}(\mathbf{X}_{\mathbf{T}}, \boldsymbol{\beta}^i)|, \quad (6)$$

$$= \frac{1}{d} \sum_{i=1}^d \log |T \mathbf{M}(\mathbf{X}, \boldsymbol{\beta}^i)|, \quad (7)$$

$$= \frac{1}{d} \sum_{i=1}^d \log \left(T^k |\mathbf{M}(\mathbf{X}, \boldsymbol{\beta}^i)| \right), \quad (8)$$

$$= \frac{1}{d} \sum_{i=1}^d \left(\log(T^k) + \log |\mathbf{M}(\mathbf{X}, \boldsymbol{\beta}^i)| \right), \quad (9)$$

$$= \frac{1}{d} \sum_{i=1}^d \left(k \log(T) + \log |\mathbf{M}(\mathbf{X}, \boldsymbol{\beta}^i)| \right), \quad (10)$$

$$= \frac{1}{d} \left(dk \log(T) + \sum_{i=1}^d \log |\mathbf{M}(\mathbf{X}, \boldsymbol{\beta}^i)| \right), \quad (11)$$

$$= k \log(T) + \tilde{\mathcal{D}}_B(\mathbf{X}). \quad (12)$$

A design $\mathbf{Z}$ is heterogeneous insofar as it consists of T different subdesigns or surveys assigned to T (groups of) respondents. It is constructed as if the choice sets of these subdesigns belong to the same large design. The heterogeneous design is thus a choice design of $T \times S$ choice sets. Because of the larger design size, $\mathbf{Z}$ allows for greater variation in the attribute levels so that more information on the parameter vector $\boldsymbol{\beta}$ can be gained. Replicated choice tasks in $\mathbf{Z}$ are allowed, but a choice task should not occur more than T times to avoid that a respondent faces it more than once. In the same vein, there are no replicated choice tasks allowed in $\mathbf{X}$. Therefore, it holds that $\tilde{\mathcal{D}}_B(\mathbf{X}_{\mathbf{T}}) \leq \tilde{\mathcal{D}}_B(\mathbf{Z})$.

To compare the estimation performance of a homogeneous design with T replications, $\mathbf{X}_{\mathbf{T}}$, to a heterogeneous design $\mathbf{Z}$ with T subdesigns, the relative $\mathcal{D}_B$ -efficiency is computed. In line with previous work on Bayesian design comparison by Kessels et al. (2011a,

2015a) and Palhazi Cuervo et al. (2016), the relative $\mathcal{D}_B$ -efficiency can be written as

$$Eff_B(\mathbf{X}_T, \mathbf{Z}) = \exp\left(\frac{\tilde{\mathcal{D}}_B(\mathbf{X}_T) - \tilde{\mathcal{D}}_B(\mathbf{Z})}{k}\right), \quad (13)$$

$$= \exp\left(\frac{k \log(T) + \tilde{\mathcal{D}}_B(\mathbf{X}) - \tilde{\mathcal{D}}_B(\mathbf{Z})}{k}\right), \quad (14)$$

$$= \exp\left(\log(T) + \frac{\tilde{\mathcal{D}}_B(\mathbf{X}) - \tilde{\mathcal{D}}_B(\mathbf{Z})}{k}\right), \quad (15)$$

$$= T \exp\left(\frac{\tilde{\mathcal{D}}_B(\mathbf{X}) - \tilde{\mathcal{D}}_B(\mathbf{Z})}{k}\right). \quad (16)$$

This expression comes down to the $\mathcal{D}_B$ -efficiency of a homogeneous design $\mathbf{X}$ relative to a heterogeneous design $\mathbf{Z}$ multiplied by T .

3 Simulation study

In this section, a simulation study is presented to identify the situations where the efficiency losses from using a single homogeneous design are small and where they are large. In the former case, the use of a homogeneous design can be justified, whereas in the latter case, a heterogeneous design should be employed.

3.1 Setup of the simulation study

The design of the simulation study is a 12-run $\mathcal{D}$ -optimal response surface design for two factors used to generate the choice design scenarios. The design appears in Table 1. The first factor is the residual degrees of freedom (DF) from using a homogeneous $\mathcal{D}_B$ -optimal design of 10 choice sets with 2 profiles to estimate a main-effects MNL model in k parameters. The second factor is the number of attributes in the $\mathcal{D}_B$ -optimal design. Both factors take values in the set $\{2, 3, 4, 5\}$. Given that the design has 10 choice sets, the number of parameters k belongs to the set $\{5, 6, 7, 8\}$. The $\mathcal{D}$ -optimal response surface design was created using the statistical software JMP, version 12 (SAS Institute, Cary, NC, USA) treating the two factors as *discrete numeric*. Such factors occupy a middle ground between completely continuous factors that can take any value on an interval and categorical factors that have a fixed number of unordered levels. Assuming the two factors as discrete numeric ensured that only the four allowable levels 2, 3, 4 and 5 for each of the factors were candidate levels for the response surface design. The 12 runs of this design describe the choice design scenarios of the simulation study. Besides the residual DF and the number of attributes, Table 1 also includes the attribute levels used in the scenarios.

<Insert Table 1 about here>

For each scenario of Table 1, a homogeneous $\mathcal{D}_B$ -optimal design of 10 choice sets with 2 profiles was constructed using JMP, version 12. This design was then evaluated as

if it was assigned to a group of T , $T = 1, 2, \dots, 10$, respondents where a heterogeneous $\mathcal{D}_B$ -optimal design of T subdesigns was taken as benchmark. For each value of T , the $\mathcal{D}_B$ -efficiency of the homogeneous design was computed relative to the heterogeneous design using (16). The setup of this simulation study allows the estimation of the full quadratic model for explaining the relative $\mathcal{D}_B$ -efficiencies in terms of the residual DF, the number of attributes and the number of subdesigns T .

Note that for the specification of the prior distribution of the Bayesian designs, all attributes were assumed ordinal the levels of which were ranked in order of preference and coded using effects-type coding. A normally distributed prior utility was chosen with mean values equally spaced between -0.6 for the first level of each attribute and 0.6 for the last level. A variance of 0.3 was specified around every prior mean value. To ensure that the variance associated with the implied utility of the last level of each attribute was also equal to 0.3 , a negative covariance was specified between the prior utilities of the attribute levels following the recommendations of Kessels et al. (2008). If L_q denotes the number of levels of attribute q , these covariances were computed using a correlation coefficient of $-1/(L_q - 1)$.

3.2 Results

3.2.1 $\mathcal{D}_B$ -efficiencies

An examination of the homogeneous and heterogeneous $\mathcal{D}_B$ -optimal designs for all 12 scenarios of Table 1 showed that none of them consist of choice sets that appear more than T times. So none of the T (groups of) respondents should face a choice set more than once. For example, scenario 5 with 5 residual DF and 2 attributes with 3 and 4 levels is the scenario with the smallest number of possible profiles (i.e. 12) and choice sets (i.e. 66). For this scenario, the heterogeneous design with $T = 10$ subdesigns consists of choice sets that appear at most 5 times. Table 2 shows the $\mathcal{D}_B$ -efficiencies from using a homogeneous design with T replications relative to a heterogeneous design with T subdesigns. Figure 1 displays the $\mathcal{D}_B$ -efficiencies for scenarios 1–8 of Table 1 involving 2 or 5 residual DF and 2, 3, 4 or 5 attributes against the number of subdesigns T .

<Insert Table 2 and Figure 1 about here>

Similar to the findings of Sándor and Wedel (2005), Table 2 and Figure 1 reveal that homogeneous designs with $T > 1$ replications are outperformed by heterogeneous designs in all scenarios. Also, beyond 3 to 4 subdesigns, the $\mathcal{D}_B$ -efficiencies remain constant. For a given number of respondents, this means that a larger number of subdesigns in a heterogeneous design does not make the parameter estimates more precise. Note that in the simulation study of Sándor and Wedel (2005), about 4 to 6 subdesigns were found optimal for a heterogeneous design. The subdesigns they generated had 12 or 18 choice sets with 2 profiles of 5 3-level attributes and 10 or 15 choice sets with 3 profiles of 4 3-level attributes.

In the same vein, this result also confirms studies by Bliemer and Rose (2011) and Rose and Bliemer (2013) on the effect of the number of choice tasks on design efficiency.

They observed that small designs are usually good enough and that up to certain number of choice tasks, there is no more benefit to adding more variety. For example, Rose and Bliemer (2013) studied a design scenario with 2 profiles per choice set and 4 3-level attributes and found that there is no need to go beyond 12 choice tasks because no more information would be gained.

In addition to confirming previous findings, more specific results were obtained from the present study. First, the more residual DF are available when using a homogeneous design, the higher the $\mathcal{D}_B$ -efficiencies. The reason is that there is more freedom to vary the attribute levels in the design matrix which carries over to a more efficient design. Second, the fewer attributes used in a homogeneous design, the higher the $\mathcal{D}_B$ -efficiencies. In most cases, a small number of attributes generates a small number of possible profiles leading to a better coverage of the profiles in the homogeneous design, which makes it more efficient. There is, however, one exception to this result, and that is scenario 8 of Table 1 corresponding to 5 residual DF and 5 2-level attributes. For this scenario, the $\mathcal{D}_B$ -efficiencies from using a homogeneous design are higher than those corresponding to 5 residual DF and 2 (only for $T = 2$), 3 and 4 attributes. This result is most likely because all 5 attributes have 2 levels only, and although 32 profiles and 496 choice sets are possible in this case, a look at the designs shows that there is hardly any level overlap in the optimal choice sets (due to the relatively weak prior information) making the number of proper choice sets for design optimization much smaller than the theoretical candidate set. Third, there is a synergistic effect between the residual DF and the number of attributes. The fewer residual DF, the larger the effect from the number of attributes on the $\mathcal{D}_B$ -efficiencies. Vice versa, the more attributes under study, the larger the effect from the residual DF on the $\mathcal{D}_B$ -efficiencies.

To summarize, when the residual DF are large and the number of attributes and attribute levels are small, with the latter being of smaller importance, the losses in $\mathcal{D}_B$ -efficiency from using a homogeneous design are small. For example, for scenarios 5–8 of Table 1, which all involve 5 residual DF, the losses range from 2% to 4%, so that using the homogeneous design is not a bad design option.

3.2.2 Quadratic model for the $\mathcal{D}_B$ -efficiencies

To quantify the effects of the residual DF, the number of attributes and the number of subdesigns T on the $\mathcal{D}_B$ -efficiencies from using a homogeneous design, the full quadratic model in these factors was estimated. Table 3 shows the final model containing only the significant effects at a 5% level. Except for the quadratic effect in the number of attributes, all effects turn out to be significant. Note that the second-order terms have been mean-centered to make the coefficients more interpretable and to reduce multicollinearity between the effects.

<Insert Table 3 about here>

The model confirms the previous findings mentioned above. The residual DF, the number of attributes and their interaction matter most for explaining the $\mathcal{D}_B$ -efficiencies. Next comes the number of subdesigns T with all linear and quadratic terms. The more

residual DF and/or the fewer attributes under study, the higher the $\mathcal{D}_B$ -efficiencies. As an illustration of the model, Figure 2 shows the effect of an increase in the residual DF from (a) 2 to (b) 5, all other factors being constant. The interaction effect between the residual DF and the number of attributes is apparent from the reduced slope of the effect of the number of attributes when increasing the residual DF from 2 to 5.

<Insert Figure 2 about here>

4 Discussion and conclusion

Using a simulation study, this paper investigates the difference in efficiency between a single homogeneous design for T respondents versus a heterogeneous design that is T times as long. Underlying the design generation is the Bayesian MNL design framework that handles nonlinear design optimization. The study results show that a homogeneous design is almost as efficient as a heterogeneous design when the residual DF of the homogeneous design are large and, to a lesser extent, when the number of attributes and attribute levels are small. A quadratic model in the residual DF, the number of attributes and the number of subdesigns T describes the $\mathcal{D}_B$ -efficiencies from using a homogeneous design relative to a heterogeneous design.

In this model for the $\mathcal{D}_B$ -efficiencies, the residual DF was chosen as a design factor that summarizes the number of choice sets, the number of profiles in a choice set and the number of parameters to estimate. It turns out, however, that the residual DF are not fully descriptive as a design size measure to predict the $\mathcal{D}_B$ -efficiencies. Also the number of choice sets of the homogeneous design plays a role as a separate factor.

For example, in the SC experiment by Verhetsel et al. (2015), a heterogeneous design of $T = 2$ subdesigns of 20 choice sets with 2 profiles was used to investigate the impact of different dimensions of accessibility on the location decision process of logistics companies. The experiment involved 6 attributes (four of 4 levels, one of 5 levels and one of 2 levels) and $k = 17$ parameters. For this scenario, the homogeneous design of 20 choice sets with 2 profiles has 3 residual DF. The $\mathcal{D}_B$ -efficiency of the homogeneous design relative to the heterogeneous design amounts to 93.29%. Using the prediction model for the $\mathcal{D}_B$ -efficiencies for 3 residual DF, 6 attributes and $T = 2$ subdesigns, results, however, in a predicted $\mathcal{D}_B$ -efficiency of 86.08% with a 95% confidence interval [85.41%, 86.75%]*. This $\mathcal{D}_B$ -efficiency value is much smaller than the observed $\mathcal{D}_B$ -efficiency of 93.29%, where the difference is mainly due to the fact that the homogeneous design of the simulation study consists of 10 choice sets only. Doubling the number of choice sets from 10 to 20 in the homogeneous design has lead to an increase in the $\mathcal{D}_B$ -efficiency of about 7%. Although the larger homogeneous design is necessary to be able to estimate the larger number of parameters, it provides more information than the homogeneous design of the simulation study when compared to the corresponding heterogeneous designs.

Because the number of choice sets in the homogeneous design has an influence on the $\mathcal{D}_B$ -efficiencies on top of the residual DF, one can view the $\mathcal{D}_B$ -efficiencies of the

---

*Although a total of 6 attributes is outside the domain of the prediction model, extra simulations confirmed that the $\mathcal{D}_B$ -efficiency for this specific scenario equals 86%.

simulation study as lower bounds for homogeneous designs with more than 10 choice sets. The larger the size of the homogeneous design, the higher the $\mathcal{D}_B$ -efficiencies, and thus the smaller the efficiency losses from using a homogeneous design. Nevertheless, if the size of the homogeneous design is small, but the residual DF are large and, to a lesser extent, the number of attributes and attribute levels are small, the efficiency losses from using a homogeneous design are negligible and its use can be justified.

A homogeneous design also benefits from the fact that there is no survey effect, which may be the case when using different surveys of a heterogeneous design. Particularly, if random blocking is applied, not all the levels of an attribute may appear in each survey which may lead to biased estimates (Lancsar and Louviere 2008, Louviere et al. 2008). However, whether one uses a homogeneous or heterogeneous design, one should never blindly trust it, but always carefully check the survey questions to ensure they cover the entire experimental region and make sense overall. The comparison using $\mathcal{D}_B$ -efficiencies is only valid in a SC setting that renders reliable and consistent choices. That is to say, the number of choice sets and choice alternatives shown to respondents should be kept manageable to prevent respondents from losing the overview and repeatedly selecting the first alternative in a choice set, for example. The attributes and attribute levels should be carefully selected and framed and, when their numbers become too large, they should be combined into partial profiles to refrain respondents from resorting to lexicographic decision rules (Kessels et al. 2011a, 2015a). These rules concern decisions based on the levels of one attribute or a small subset of the attributes. Also, partial profiles may prevent respondents from making decisions based on the levels of the first or last attribute in a choice set. Caussade et al. (2005) shed further details on these design features that have the potential to influence the responses. Drawing on their work, I emphasize that short well-defined and statistically efficient design surveys should be constructed that keep respondents motivated and responsive, but at the same time, provide rich datasets for complex problems. This should safeguard the overall quality of the design, which depends on both statistical and response qualities (Louviere et al. 2010, Reed Johnson et al. 2013).

Finally, this comparison study of homogeneous versus heterogeneous designs has been carried out under the Bayesian MNL design framework. Sándor and Wedel (2005) also compared homogeneous versus heterogeneous locally optimal designs for the cross-sectional mixed logit model that accounts for respondent heterogeneity by assuming random parameters. The design comparison results are largely similar for both models, but for the mixed logit model, Sándor and Wedel (2005) observed larger efficiency gains for heterogeneous designs. The reason is that these designs might be more capable of capturing differences in the parameters among respondents. A related finding was observed by Rose and Bliemer (2013), who studied the effect of the number of choice tasks on the efficiency of locally optimal designs for the panel mixed logit model. The panel version of the mixed logit model takes into account that choice observations from the same respondent are correlated and is thus more realistic than its cross-sectional counterpart. Rose and Bliemer (2013) observed that especially the standard deviation parameters of the panel mixed logit model are more precisely estimated when using larger designs. Because the number of choice tasks has to be manageable for the respondents, there is much value in

constructing heterogeneous designs for the panel mixed logit model. This would require the development of a proper blocking procedure to account for repeated choices from the same respondent. Such heterogeneous design procedure could then be compared to the individually adapted sequential Bayesian design approach by Yu et al. (2011) for generating panel mixed logit designs for individual respondents. Also the new heterogeneous design construction approach by Liu and Tang (2015) could be adapted and evaluated for generating these designs. Developing practical design approaches for the panel mixed logit model is a promising area for future research.

Acknowledgement

Roselinde Kessels is grateful to the Fund for Scientific Research Flanders (FWO) for her postdoctoral research grant. She thanks Stephane Hess and two anonymous referees for their valuable comments which have significantly improved the paper.

References

- Bliemer, M. C. J. and Rose, J.M. (2010). Construction of experimental designs for mixed logit models allowing for correlation across choice observations, *Transportation Research Part B* **44**: 720–734.
- Bliemer, M.C.J. and Rose, J.M. (2011). Experimental design influences on stated choice outputs: An empirical study in air travel choice, *Transportation Research Part A: Policy and Practice* **45**: 63–79.
- Caussade, S., Ortúzar, J. de D., Rizzi, L.I., Hensher, D.A. (2005). Assessing the influence of design dimensions on stated choice experiment estimates, *Transportation Research Part B: Methodological* **39**: 621–640.
- ChoiceMetrics (2014). *Ngene 1.1.2 User Manual & Reference Guide*, Australia.
- Gotwalt, C.M. (2010). Addendum to “Fast computation of designs robust to parameter uncertainty for nonlinear settings”, *Technometrics* **52**: 137.
- Gotwalt, C.M., Jones, B.A. and Steinberg, D.M. (2009). Fast computation of designs robust to parameter uncertainty for nonlinear settings, *Technometrics* **51**: 88–105.
- Hess, S. and Rose, J.M. (2009). Should reference alternatives in pivot design SC surveys be treated differently?, *Environmental and Resource Economics* **42**: 297–317.
- Kessels, R., Jones, B. and Goos, P. (2011a). Bayesian optimal designs for discrete choice experiments with partial profiles, *Journal of Choice Modelling* **4**: 52–74.
- Kessels, R., Jones, B. and Goos, P. (2015a). An improved two-stage variance balance approach for constructing partial profile designs for discrete choice experiments, *Applied Stochastic Models in Business and Industry* **31**: 626–648.
- Kessels, R., Jones, B., Goos, P. and Vandebroek, M. (2008). Recommendations on the use of Bayesian optimal designs for choice experiments, *Quality and Reliability Engineering International* **24**: 737–744.

- Kessels, R., Jones, B., Goos, P. and Vandebroek, M. (2011b). The usefulness of Bayesian optimal designs for discrete choice experiments, *Applied Stochastic Models in Business and Industry* **27**: 173–188.
- Kessels, R., Jones, B., Goos, P. and Vandebroek, M. (2011c). Rejoinder: The usefulness of Bayesian optimal designs for discrete choice experiments, *Applied Stochastic Models in Business and Industry* **27**: 197–203.
- Kessels, R., Van Herck, P., Dancet, E., Annemans, L. and Sermeus, W. (2015b). How to reform western care payment systems according to physicians, policy makers, health-care executives and researchers: A discrete choice experiment, *BMC Health Services Research* **15**: 191.
- Kupfer, F., Kessels, R., Goos, P., Van de Voorde, E. and Verhetsel, A. (2016). The origin–destination airport choice for all-cargo aircraft operations in Europe, *Transportation Research Part E* **87**: 53–74.
- Lancsar, E. and Louviere, J. (2008). Conducting discrete choice experiments to inform healthcare decision making: A user’s guide, *PharmacoEconomics* **26**: 661–677.
- Liu, Q. and Tang, E. (2015). Construction of efficient heterogeneous choice designs: A new approach, *Marketing Science* **34**: 346–366.
- Louviere, J.J., Pihlens, D. and Carson, R. (2010). Design of discrete choice experiments: A discussion of issues that matter in future applied research, *Journal of Choice Modelling* **4**: 1–8.
- Louviere, J.J., Street, D., Burgess, L., Wasi, N., Islam, T. and Marley, A.A.J. (2008). Modelling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information, *Journal of Choice Modelling* **1**: 128–163.
- Luyten, J., Kessels, R., Goos, P. and Beutels, P. (2015). Public preferences for prioritizing preventive and curative health care interventions: A discrete choice experiment, *Value in Health* **18**: 224–233.
- Palhazı Cuervo, D., Kessels, R., Goos, P. and Sørensen, K. (2016). An integrated algorithm for the optimal design of stated choice experiments with partial profiles, *Transportation Research Part B: Methodological* **93**: 648–669.
- Reed Johnson, F., Lancsar, E., Marshall, D., Kilambi, V., Mühlbacher, A., Regier, D.A., Bresnahan, B.W., Kanninen, B. and Bridges J.F.P. (2013). Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force, *Value in Health* **16**: 3–13.
- Rose (2011). Discussion on “The usefulness of Bayesian optimal designs for discrete choice experiments”, *Applied Stochastic Models in Business and Industry* **27**: 193–196.
- Rose, J.M. and Bliemer, M.C.J. (2013). Sample size requirements for stated choice experiments, *Transportation* **40**: 1021–1041.
- Rose, J.M., Bliemer, M.C.J, Hensher, D.A. and Collins, A.T. (2008). Designing efficient stated choice experiments in the presence of reference alternatives, *Transportation Research Part B: Methodological* **42**: 395–406.

- Sándor, Z. and Wedel, M. (2001). Designing conjoint choice experiments using managers' prior beliefs, *Journal of Marketing Research* **38**: 430–444.
- Sándor, Z. and Wedel, M. (2002). Profile construction in experimental choice designs for mixed logit models, *Marketing Science* **21**: 455–475.
- Sándor, Z. and Wedel, M. (2005). Heterogeneous conjoint choice designs, *Journal of Marketing Research* **42**: 210–218.
- Verhetsel, A., Kessels, R., Goos, P., Zijlstra, T., Blomme, N. and Cant, J. (2015). Location of logistics companies: A stated preference study to disentangle the impact of accessibility, *Journal of Transport Geography* **42**: 110–121.
- Verhetsel, A., Kessels, R., Zijlstra, T. and Van Bavel, M. (2016). Housing preferences among students: collective housing versus individual accommodations? A stated preference study in Antwerp (Belgium), *Journal of Housing and the Built Environment*, online available at <http://dx.doi.org/10.1007/s10901-016-9522-5>.
- Yu, J., Goos, P. and Vandebroek, M. (2009). Efficient conjoint choice designs in the presence of respondent heterogeneity, *Marketing Science* **28**: 122–135.
- Yu, J., Goos, P. and Vandebroek, M. (2011). Individually adapted sequential Bayesian conjoint-choice designs in the presence of consumer heterogeneity, *International Journal of Research in Marketing* **28**: 378–388.

Tables and figures

Table 1: Twelve design scenarios used in the simulation study.

Scenario	Residual degrees of freedom (DF)	Number of attributes	Attribute levels
1	2	2	5,5
2	2	3	3,4,4
3	2	4	2,3,3,4
4	2	5	2,2,3,3,3
5	5	2	3,4
6	5	3	2,3,3
7	5	4	2,2,2,3
8	5	5	2,2,2,2,2
9	3	3	3,3,4
10	3	5	2,2,2,3,3
11	4	2	4,4
12	4	4	2,2,3,3

Table 2: $\mathcal{D}_B$ -efficiencies of assigning a homogeneous (10 choice set) design for each scenario of Table 1 (in terms of Residual DF and Number of attributes) to T respondents relative to using a heterogeneous design with T subdesigns.

T	Scenario											
	2-2	2-3	2-4	2-5	5-2	5-3	5-4	5-5	3-3	3-5	4-2	4-4
1	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
2	95.40	93.37	90.00	86.16	98.18	96.76	96.68	98.51	94.39	88.42	97.27	94.64
3	94.37	91.98	88.04	83.65	97.85	96.21	96.21	97.41	93.25	86.64	96.64	93.70
4	93.95	91.41	87.29	82.63	97.76	96.22	95.91	96.96	92.80	85.99	96.44	93.38
5	93.78	91.10	86.90	82.20	97.70	96.12	95.83	96.85	92.58	85.71	96.35	93.17
6	93.66	90.90	86.68	81.95	97.67	96.08	95.77	96.83	92.46	85.56	96.33	93.08
7	93.59	90.83	86.51	81.80	97.65	96.06	95.73	96.78	92.38	85.48	96.30	93.03
8	93.54	90.73	86.41	81.66	97.65	96.04	95.70	96.76	92.33	85.39	96.28	93.00
9	93.48	90.66	86.35	81.58	97.64	96.03	95.69	96.73	92.30	85.34	96.26	92.98
10	93.47	90.63	86.30	81.53	97.64	96.02	95.67	96.71	92.27	85.31	96.26	92.96

Table 3: Final quadratic model for the $\mathcal{D}_B$ -efficiencies of Table 1 in the residual DF, the number of attributes and the number of subdesigns T using mean-centered values. The model effects are significantly different from zero at a 5% level.

Effect	Estimate	Std Error	t Ratio	p Value
Intercept	91.5045	0.3591	254.81	< 0.0001
Residual DF	2.6969	0.0529	50.97	< 0.0001
Nr Attributes	-2.1980	0.0596	-36.91	< 0.0001
(Residual DF-3.5) $\times$ (Nr Attributes-3.5)	1.2207	0.0475	25.70	< 0.0001
T	-0.2020	0.0257	-7.85	< 0.0001
$(T-6) \times (T-6)$	0.0540	0.0114	4.76	< 0.0001
$(T-6) \times$ (Residual DF-3.5)	0.0749	0.0205	3.66	0.0004
$(T-6) \times$ (Nr Attributes-3.5)	-0.0554	0.0231	-2.40	0.0182
(Residual DF-3.5) $\times$ (Residual DF-3.5)	0.1469	0.0707	2.08	0.0404

Figure 1: $\mathcal{D}_B$ -efficiencies of assigning a homogeneous (10 choice set) design for scenarios 1–8 of Table 1 (in terms of Residual DF and Number of attributes) to T respondents relative to using a heterogeneous design with T subdesigns.

Figure 2: Model predictions for the $\mathcal{D}_B$ -efficiency of using a homogeneous (10 choice set) design showing the effect of an increase in the Residual DF from (a) 2 to (b) 5 given $T = 6$ and Number of attributes = 5.