Erasmus University Rotterdam

Erasmus School of Economics

MASTER'S THESIS BEHAVIORAL ECONOMICS

Extent of Irrationality of the Consumer: Combining the Critical Cost Efficiency and Hautman Maks Indices

Author: **Matěj Opatrný**

Student ID Number: **449297**

Supervisor: **dr. Jan Heufer**

Second Assesor: **Paul van Bruggen**

Date of Final Version: **May 23, 2017**

Abstract

In this thesis, we use the Critical Cost Efficiency Index and the Hautman Maks Index to evaluate the consistency of subjects in the dataset. As a result, we show that by simply allowing subjects for one significant mistake (by removing the worst observation with the highest wasted budget), the consistency of the dataset increases by 6 percentage points. Furthermore, we demonstrate that by excluding the worst observation per subject, the fraction of subjects wasting 5% or less of their budget increases from 45% to 64%. Therefore, the larger and more consistent dataset can be used for further study. Finally, we apply the aforementioned findings to various socio–economic groups. The results indicate that the highest improvement in terms of the efficiently spent budget can be seen among retired and 65+ aged subjects.

JEL Classification D12, D14, D81, D83, D91, G11
Keywords Critical Cost Efficiency Index, Hautman Maks Index, Socio-economic groups, observation

Author's e-mail matej.opatrny@gmail.com
Supervisor's e-mail heufer@ese.eur.nl

Contents

List of Tables

List of Figures

1 Introduction

Recently, a great deal of emphasis has been placed on understanding a consumer behaviour and establishing a pattern that is able to predict consumer choices or at least to measure the extent of inconsistency of choices (van Bruggen & Heufer, 2017).[1][2] The classical utility maximization model has traditionally been used as a starting point for the measure of consumer' consistency. As a consequence, Afriat (1972) tries to measure the level of inconsistency from the classical model. More precisely, Afriat (1972) introduces an index which captures the *fraction of inconsistently used budget* at the level of an individual consumer. Furthermore, Houtman & Maks (1985) contribute to this topic with the idea to measure the maximum *number of choices* consistent with the classical model. Using the CentERpanel data, this thesis puts the aforementioned techniques together and shows the application and relationship between both approaches. We argue that by excluding the worst observation per subject, one will obtain the higher level of consistency of the dataset accompanied with the larger fraction of subjects satisfying this level. This implicates that one can test the theoretical models with the larger and more consistent dataset. Therefore, this study shows that excluding the worst observation per subject may be better approach than excluding 'inconsistent' subjects to achieve larger and more consistent dataset. The dataset consists of 25 budget allocation decision choices per subject in a two–dimensional commodity space.

Our sample indicates that, on average, subjects waste around 12% of their budget. This corresponds with almost 3 inconsistent choices in the sense of using the whole budget efficiently. We show that by simply allowing subjects for one significant mistake (in terms of removing the worst observation with the highest wasted budget), the consistency of the dataset increases by 6 percentage points (pp). The other important advantage of excluding the worst observation per subject is that a larger number of subjects satisfy a higher threshold of consistently used budget (the threshold depends on the subjective opinion of the researcher).[3] Using our sample, we demonstrate that by excluding the worst observation per subject, the fraction of subjects wasting 5% or less of their budget increases from 45% to 64%. Therefore, a larger and more consistent dataset can be used for further analysis by removing only a single choice from each data set. Furthermore, using Bronars (1987) and Beatty & Crawford (2011) approaches for measuring the demand of the theoretical restriction placed on the data, we show that excluding the worst choice per subject yields higher optimal level of CCEI Index (resulting in wasting less budget) with larger fraction of subjects satisfying this optimal level of CCEI Index. Finally, we apply the aforementioned findings to various socio–economic groups. The results indicates that the highest improvement in terms of efficiently spent budget can be seen among retired and 65+ aged subjects.

This thesis is structured as follows. In section 2, we summarize the core literature that is at the base of this empirical study. In section 3, we describe the dataset and the methodological background for this analysis. In section 4, we provide the results of this study. The conclusion can be found in section 5.

[1] By consistency is meant that choices do satisfy Generalized Axiom of Revealed Preferences (GARP), described in section 3.

[2] For example (Sippel, 1997), (Harbaugh & Krause, 2000), (Mattei, 2000), (Andreoni & Miller, 2002), (Février & Visser, 2004), (Fisman *et al.*, 2007), (Dickinson, 2009), (Banerjee & Murphy, 2011), (Camille *et al.*, 2011), (Dawes *et al.*, 2011), (Visser & Roelofs, 2011), (Bruyneel *et al.*, 2012), (Becker *et al.*, 2013), (Burghart *et al.*, 2013), and (Ahn *et al.*, 2014).

[3] Varian (1991) in his study suggest that to waste 5% or less of the budget is a reasonably close to utility maximization.

2 Related Literature

The literature on testing the consumer revealed preferences is rich and includes a variety of classical papers and recent contributions. Since Tversky & Kahneman (1975) demonstrated that consumer decision making is influenced by heuristics and biases, such as anchoring, framing, representativeness, loss aversion, etc., the demand for a more sophisticated method of measuring the consumer utility function has increased. However, as Choi *et al.* (2007) state, the behaviour of subjects is very complex and it is almost impossible to classify it in a simple taxonomy. Moreover, they point out that most subjects behave as utility maximizers.[4] Therefore, to measure the extent of deviation from the utility maximization model, we employ Afriat's (1967) approach, which states that a finite number of individuals' choices from a series of budget sets can be described by a well–behaved (monotonic, continuous, and concave) utility function if and only if subjects satisfy a condition he called cyclic consistency. This statement was refined and later proved by Varian (1982), who shows that satisfying the Generalized Axiom of Revealed Preferences (GARP) is a necessary and sufficient condition for consumer choices to be consistent with the maximization of a continuous, concave, locally nonsatiated, and weakly monotonic utility function.

Furthermore, Afriat (1972) introduced the Critical Cost Efficiency Index (CCEI), which was later used to measure the fraction by which all budgets need to be shifted to satisfy GARP. In a study by Varian (1991), authors apply the measure of the CCEI Index for 38 subjects (the dataset was collected by Battalio *et al.* (1973) and consists of 38 long–term patients operating in a token economy – they can exchange tokens for goods such as cigarettes, etc. – at the Central Islip State Hospital) and find that their choices are very close to optimal behaviour (maximization of utility). In a study by Choi *et al.* (2014), the authors use the CCEI Index to measure the extent of irrationality across different socio–economic groups. They find that, for example, retired subjects have the highest level of wasteful budget allocation (almost 17%), whereas young subjects are the best utility maximizers, wasting 'only' around 8% of their budget. Another study by Harbaugh *et al.* (2001) shows that the level of inconsistency does not substantially differ between children and adults, and therefore, the same modelling of choice behaviour can be used for both groups. Another study using the CCEI Index done by Andreoni & Miller (2002) shows that altruistic behavior can be consistent with the GARP axiom and it can therefore be considered 'rational'.

A different approach that measures the extent of irrationality was introduced by Houtman & Maks (1985). In their study, they introduce the Hautman Maks (HM) Index, which measures the maximum number of observations satisfying the Strong Axiom of Revealed Preference (SARP) as well as an algorithm which computes the HM Index. Based on the Houtman & Maks (1985) approach, Gross & Kaiser (1996) construct an algorithm that computes the maximal subset consistent with the Weak Axiom of Revealed Preferences (WARP) for any dimensional cases, while demonstrating its application on experimental choice data. Furthermore, Heufer & Hjertstrand (2015) apply Gross & Kaiser (1996) algorithm to find the maximal subset consistent with the WARP, making use of Banerjee & Murphy (2006) result that shows that Weak Generalized Axiom of Revealed Preferences (WGARP) and the GARP are equivalent in the two-dimensional commodity space. Moreover, they use a new Mixed Integer Linear Programming (MILP) approach for higher dimensional commodity space.

[4] In terms of maximizing a complete, transitive preferences ordering over some portfolios (Choi *et al.*, 2007).

Another approach to analyse the choices made by individual consumers has been used by Bronars (1987). In their study, besides the other methods, randomly generated choices are compared to the actual choices to check the power of the data. Bronars (1987) method has been further developed by Beatty & Crawford (2011), in which they challenge the nature of restrictions that fundamental economic theory places on data. When they account for a quite undemanding nature of the restrictions imposed on the data, the performance of the fundamental model is far less impressive. They argue that using their sample (data are from Spanish Continuous Family Expenditure Survey and consists of 21,866 observations), the economic model outperforms randomly generated data only by 4.5% in terms of satisfying the restrictions of the model, while taking into account the power of the restrictions.

3 Data and Methodology

The thesis uses the CentERpanel dataset with a sample of over 2,000 households and 5,000 participants from the Netherlands (Choi *et al.*, 2014). Respondents answered an online survey, where, besides providing answers about the experiment, they indicated individual demographic and economic information. The design of the experimental questions was made by Choi *et al.* (2014) and is as follows. Subjects made 25 decision choices in total, in a two–dimensional budget space. Each choice represents the allocation between accounts x (horizontal line) and y (vertical line). The actual payoffs were determined according to the subject's choice; the subject received the points allocated to one of the accounts x or y, which were chosen at random and equally likely (Choi *et al.*, 2014). In total, the sample consists of 1,372 respondents, however, only 1,182 subjects fully completed the survey. Table 1 shows a descriptive summary of the dataset.

Table 1: Descriptive Statistics

	Completed the Survey		
	No - Dropouts	**Yes**	**Total**
	Col %	Col %	Col %
Gender			
Male	62.1	54.6	55.6
Female	37.9	45.4	44.4
Age			
16-34	3.2	18.5	16.4
35-49	12.1	26.1	24.2
50-64	38.4	35.6	36.0
65+	46.3	19.7	23.4
Education			
High	34.7	36.5	36.3
Low	42.6	33.7	34.9
Medium	22.6	29.8	28.8
Income			
0-2,5k	40.0	22.8	25.1
2,5k-3,49k	22.1	25.5	25.1
3,5k-4,99k	15.8	29.2	27.3
5k+	22.1	22.5	22.4
Occupation			
House	7.9	11.6	11.1
Others	10.0	14.4	13.8
Paid	39.5	53.1	51.2
Retired	42.6	20.9	23.9
Partner			
No	32.1	19.1	20.9
Yes	67.9	80.9	79.1
N	190	1,182	1,372

Source: CentERpanel Data

Let us define terms used in the thesis. The commodity space is R^L_+ and the price comes from the R^L_{++} space, where $L \geq 2$ means the number of commodities (Heufer & Hjertstrand, 2015). However, in our case, subjects choose from a two–dimensional budget space, therefore, $L = 2$. The subject's budget set is defined as follows: $B = B(\mathbf{p}) = \{x \in R^2_+ : \mathbf{px} \leq 1\}$, where $\mathbf{p} = (p_1, p_2)' \in R^2_{++}$ is the price vector and income is normalised to 1 (Heufer & Hjertstrand, 2015). Therefore, we observe N budgets, in our case 25 per subject, and the decision choices made by the particular subject. Moreover, observations are written with a subscript such that bundle $\mathbf{x}^i$ is the observed choice for budget $B(\mathbf{p}^i)$, assuming that $\mathbf{x}^i\mathbf{p}^i = 1$. Thanks to the fact that price vectors characterise budgets we can write the entire set of N observation as $\{x^i, p^i\}^N_{i=1}$ (Heufer & Hjertstrand, 2015).

As was mentioned in section 2, Varian (1982) proved that satisfying the Generalized Axiom of Revealed Preference (GARP):

Definition 1. Let define $x^i R^0 x$ if $p^i x^i \geq p^i x$ and $x^i P^0 x$ if $p^i x^i > p^i x$. Let R be a transitive closure R^0, meaning, that there exists a sequence $x^j, ..., x^k$, such that $x^i R^0 x^j R^0 ... x^k R^0 x$. Let define $x^i P x$ if $x^i R x^j P^0 x^k R x$ (Heufer & Hjertstrand, 2015).

Varian (1982): Set of observations $\{x^i, p^i\}^N_{i=1}$ satisfies the GARP if for all $i, j = 1, ..., N$ it holds that not $x^i P^0 x^j$ whenever $x^j R x^i$.

is a necessary and sufficient condition to maximize continuous, locally nonsatiated, concave, and weakly monotonic utility function. Varian (1982) thoughts are based on the following Afriat's theorem (Afriat, 1967) pp. 946:

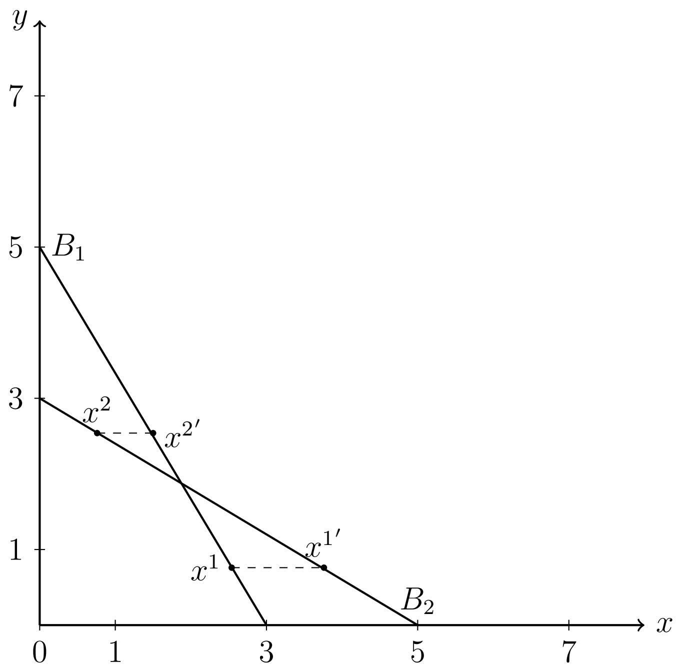
Theorem 1. (Afriat, 1967) The following conditions are equivalent:

1. There exists a nonsatiated utility function that rationalizes the data.
2. The data satisfies "cyclical consistency"; that is, $p^r x^r \geqq p^r x^s, p^s x^s \geqq p^s x^t, \ldots, p^q x^q \geqq p^q x^r$ implies $p^r x^r = p^r x^s, p^s x^s = p^s x^t, \ldots, p^q x^q = p^q x^r$.
3. There exist numbers $U_i, \lambda^i > 0$, $i = 1, \ldots, n$, such that
   $U^i \leqq U^j + \lambda^j p^j (x^i - x^j)$ for $i, j = 1, \ldots, n$.
4. There exists a nonsatiated, continous, concave, monotonic function that rationalizes the data.

In the study by Varian (1982), Varian shows that 'cyclical consistency' is equal to GARP. As a consequence of Varian's findings, if revealed preferences satisfy the GARP, standard economic models can be applied to analyse a subject's behavior (Harbaugh *et al.*, 2001).

Figure 1 shows the intuition behind the violation of the GARP axiom. If the subject choose x^1 on the budget B_1 , then x^1 is revealed preferred to $x^{2'}$. Assuming monotonic utility function, the utility of bundle $x^{2'}$ is higher than the utility of bundle x^2 (it has more of at least one good). Therefore, the subject revealed prefers x^1 to x^2. A similar argument would also imply that x^2 is revealed preferred to x^1. In conclusion, this contradiction yields that choices x^1 and x^2 cannot be the result of a rational choice, and, as a consequence, choices x^1 and x^2 do violate GARP. Put differently, the GARP requires that if a subject chooses x^1 in the first round with budget B_1 then a subject cannot choose x^2 in the second round with budget B_2, when any alternative choice with at least as much good as in x^1, and more of at least one, is available – e.g. $x^{1'}$.

Figure 1: Ilustration of GARP Violation



Nevertheless, GARP provides only two outcomes: either the data satisfies or does not satisfy the conditions for GARP. Therefore, Afriat (1972) came with a Critical Cost Efficiency Index (CCEI), which measures the fraction by which all the budgets have to be shifted to satisfy GARP. In other words, Afriat (1972) developed a tool to measure the extent of violation of GARP. Put precisely, let for any $e \in [0, 1]$

$$x^i R^0(e) x^j \Leftrightarrow e p^i x^i \geq p^i x^j,$$

and let $R(e)$ be the transitive closure of $R^0(e)$. Moreover, let $e^{'}$ be the largest value for which the relation $R(e^{'})$ satisfies GARP. The number $e^{'}$ is defined as the CCEI Index associated with a particular subject (Choi *et al.*, 2014). Let GARP(e) be the relaxed version of GARP:

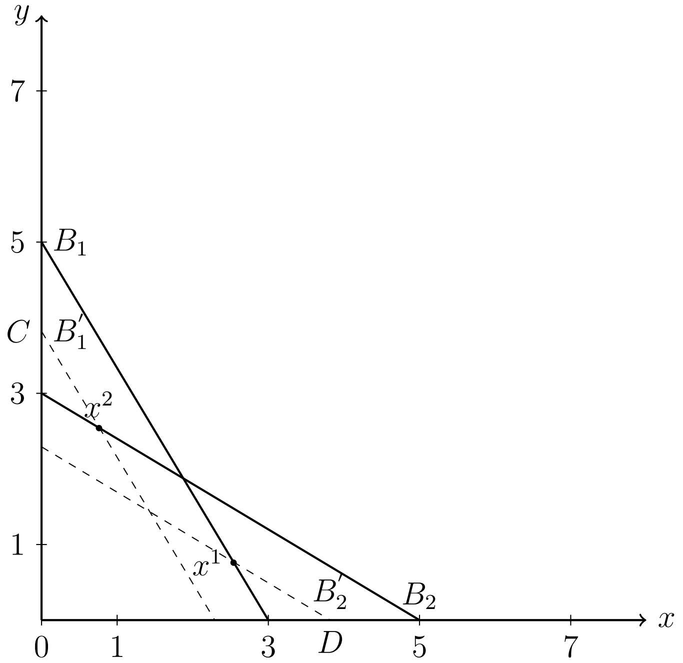
Definition 2. Varian (1991) A set of observation $\{x^i, p^i\}_{i=1}^N$ satisfies GARP(e) for some $e \in [0, 1]^N$ if for all $i, j = 1, ..., N$, it holds that not $x^i P^0(e^i) x^j$ whenever $x^j R(e^j) x^i$.

If $e = 1$ we have standard version of GARP and if $e = 0$ all observations satisfy a GARP test. Therefore, the convenient form of measurement is to see how close e is to 1, while still satisfying GARP(e) (Varian, 1991). We employ the bisectional method to determine the CCEI Index.[5]

Figure 2 illustrates the shifting budgets from figure 1 B_1, B_2 in order to satisfy GARP(e). As a result, the choices x^1, x^2 on the new budget lines $B_2^{'}$ and $B_1^{'}$ do not violate GARP(e). In other words, to remove all violation of GARP one would have to lower the budget B_1 by $C/5$ or the budget B_2 by $D/5$ (CCEI Index in this case is either $C/5$ or $D/5$, whatever is the higher value).

[5] The intuition behind the bisectional method is as follows: we set $e_0 = 1/2$ and see whether GARP(e) is satisfied, if not, we set $e_1 = 1/2e_0$ and so on. Similarly, if GARP(e) satisfies $e_0 = 1/2$, we go the other direction and set $e_1 = 1 - 1/2e_0$ and check whether GARP(e) is satisfied. We repeat this process 10 times. See Burden & Faires (1985) for details about the algorithm.

Figure 2: Illustration of Intuition behind CCEI Index



Houtman & Maks (1985) introduce another feasible method to measure the violation of GARP – the Hautman Maks Index (HM Index). The HM Index is the maximum number of observations that satisfy GARP. More specifically, let $v = (v^1, ..., v^N)$ be a vector that takes values equal to 1 or 0 for all $i = 1, ..., N$. Let define the relation $R^0(v)$ as $x^i R^0(v^i) x^j$ if $v^i p^i x^i \geq p^i x^j$, let $R(v)$ be the transitive closure of $R^0(v)$, and let $P^0(v^i)$ if $v^i p^i x^i > p^i x^j$ (Heufer & Hjertstrand, 2015). Furthermore, as Heufer & Hjertstrand (2015) pp.88 states:

Definition 3. (Heufer & Hjertstrand, 2015) A set of observations $\{x^i, p^i\}_{i=1}^N$ satisfies GARP(v) for some $v \in \{0,1\}^N$ if for all $i, j = 1, ..., N$, it holds that not $x^i P^0(v^i) x^j$ whenever $x^j R(v^j) x^i$.

Therefore, the solution to the following maximization problem such that GARP(v) holds and $v \in \{0,1\}^N$ is the HM Index (Heufer & Hjertstrand, 2015) :

$$HM = \underbrace{max}_{\text{v}} \sum_{i=1}^{N} \frac{v^i}{N}$$

However, the optimal threshold of the CCEI Index or the HM Index remains unclear.[6] To shed more light on this issue Beatty & Crawford (2011) provide a way, which could determine the optimal threshold for the CCEI and HM indices. The Beatty & Crawford (2011) approach consists of two main parts. The first one computes the pass rate of observations, denoted by $r \in [0, 1]$ (r is equal to one if the data satisfies the revealed preference restrictions, and zero if it misses by the maximum possible amount (Beatty & Crawford, 2011)). The second one determines how demanding the theoretical restrictions posed on the subject's choices are, denoted by $a \in [0, 1]$ (in

[6]For example, in Choi *et al.* (2007) study, the authors used the CCEI threshold of 0.80, which is based mainly on author's subjective opinion.

the sense of the fraction of all possible choice combination that satisfies the restrictions). As a result, Beatty & Crawford (2011) combine these two parts into function $m(r, a)$, which satisfies **monotonicity** – $m(0, 1) > m(1, 0)$ – (the model satisfying more demanding restrictions is better than the one satisfying less demanding restrictions), **equivalence** – $m(0, 0) = m(1, 1)$ – (a situation when no restriction are placed on the data is equal to the situation when no data is ruled out), and **aggregability** – $m(\lambda r_1 + (1-\lambda)r_2, \lambda a_1 + (1-\lambda)a_2) = \lambda m(r_1, a_1) + (1-\lambda)m(r_2, a_2)$ – (the measure is additive over heterogeneous subjects, therefore, sample average results can be calculated). Given the aforementioned axioms and using the following Selten's theorem:

Theorem 2. (Selten, 1991) The function $m = r - a$ satisfies monotonicity, equivalence, and aggregability. If the function $\tilde{m}(r, a)$ also satisfies these axioms, then there exist real numbers $\beta, \gamma > 0$ such that $\tilde{m}(r, a) = \beta + \gamma m$.

Beatty & Crawford (2011) explain that not only does the simple difference measure $(r - a)$ satisfy these axioms, but any measures satisfying these axioms are positive linear transformations of this difference. Furthermore, the resulting $m \in [-1, 1]$ can be interpreted as a pass/fail rate indicator taking into account the ability to find the rejections (Beatty & Crawford, 2011). Therefore, as m approaches minus one, the restrictions are so flexible that anyone can pass them. As a result, the data has no inference value. As m approaches one, we have extremely demanding restrictions accompanied with the data satisfying them, which is the sign of quantitatively successful model (Beatty & Crawford, 2011). As m approach zero, the data simply mirrors the probability of passing the restrictions given a uniform distribution over all possible choices. Another explanation of $m \approx 0$, so that the data perform as well as a uniformly generated data, is provided in a Bronars (1987) study. In his study, Bronars (1987) conducted a statistical power test by measuring Pr(Rejecting $H_0 \mid H_0$ is false), where the H_0 hypothesis is 'optimizing behavior' and the alternative hypothesis is 'uniform random choices over the outcome space' (Beatty & Crawford, 2011). Putting the aforementioned facts together, we establish the measure of the power of the data, $\hat{m}_{CCEI} = \hat{r} - \hat{a}$, where $\hat{m}_{CCEI}$ indicates a particular value of the power on the certain CCEI Index level. As such, $\hat{a}$ stands for the fraction of subjects with randomly generated observations satisfying GARP, and $\hat{r}$ stands for fraction of subjects with actual observations satisfying GARP. Therefore, we identify the CCEI Index that maximizes the $\hat{m}_{CCEI}(\hat{r}, \hat{a})$.[7] The implication is that we find the optimum fraction of the subjects satisfying the optimal CCEI level given our data.

In addition, using the CentERpanel data, we establish the relationship between CCEI and HM indices to shed more light on the issue of the data efficiency. We use the Gross & Kaiser (1996) algorithm to calculate the HM Index.[8] We apply the algorithm to a two–dimensional case, which allows us to use the Weak Generalized Axiom of Revealed Preferences (WGARP):

Definition 4. (Banerjee & Murphy, 2006) A set of observation $\{x^i, p^i\}_{i=1}^N$ satisfies WGARP if for all $i, j = 1, ..., N$, it holds that not $x^i P^0 x^j$ whenever $x^j R^0 x^i$.

Thanks to Banerjee & Murphy (2006) proof that in two–dimensional case the GARP and WGARP are equivalent, we use the WGARP in the algorithm, which remarkably simplifies it.

[7]In next section 4, figures 7, 8, and 9 show the value of $\hat{m}_{CCEI}$ as a maximum distance between the line indicating randomly generated observations and the line indicating actual observations.

[8]The link for the Matlab code is provided in the Appendix.

As Heufer & Hjertstrand (2015) states, Houtman & Maks (1985) and Gross & Kaiser (1996) use a graph–theoretic approach. Therefore, each observation can be interpreted as the node of a graph. When observations i and j violate WGARP, then these two observations are adjacent (Heufer & Hjertstrand, 2015). Consequently, the number of adjacent nodes is defined as the degree of a node, degr(i). Let A_i be the set of nodes adjacent to node i, and let $1A_i$ be the set of nodes adjacent to node i with degree 1. The algorithm works as follows Heufer & Hjertstrand (2015) pp.88:

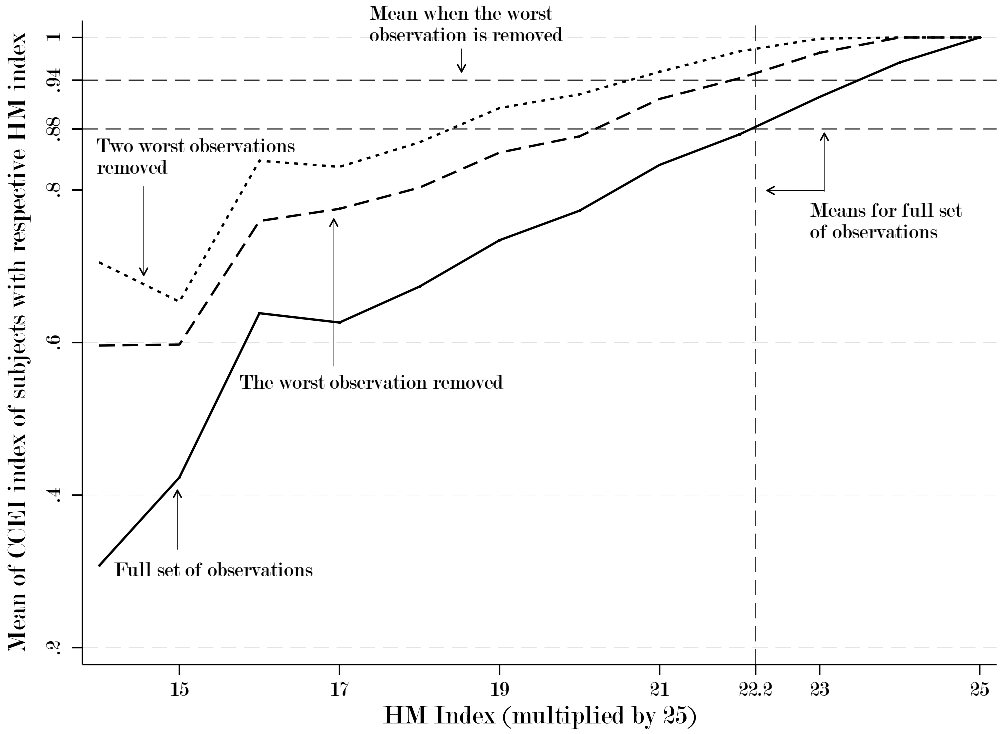
> The algorithm consists of two parts. First, whenever degr(i)= $max_{j\in 1,\dots,N}$ degr(j) and degr(k)<degr(i) for all $k \in A_i$, remove i. Repeat this step until no index is removed anymore. Second, whenever degr(i) = degr(h) = $max_{j\in 1,\dots,N}$ degr(j) and $h \in A_i$, then (1) if $1A_i \neq \emptyset$, remove i, (2) if $1A_h \neq \emptyset$, remove h, (3) if $1A_i = 1A_h = \emptyset$, remove either i or h. Again, repeat this step until no index is removed anymore. All nodes not removed in this process belong to the set of indices consistent with Warp. Gross & Kaiser (1996) point out that there is a special case in which the algorithm will fail to provide a maximal subset. However, they argue that this case is extremely rare, and in any case, the algorithm provides a lower bound.

The main contribution of this study is to show how to increase data efficiency, and therefore, the power of inference made by the data. Rather than excluding inconsistent subjects from the dataset, we show that removing the worst observation per subject leads to the fact that a larger fraction of subjects satisfy a higher threshold of the CCEI Index. We proceed as follows: we combine the Gross & Kaiser (1996) and Houtman & Maks (1985) algorithm with computations of the CCEI Index to determine the two worst observations per subject. The worst observation receives the minimum CCEI Index, while, the second worst observation obtains the second minimum CCEI Index. This allows us to determine the highest CCEI Index per subject in both scenarios, when we remove only the worst observation and when we exclude the two worst observations. We apply the aforementioned approach to various socio–economic groups to demonstrate the positive effect of removing the worst observation per subject (see section 4, figure 6 for details). Moreover, using the Bronars (1987) and the Beatty & Crawford (2011) approach, we conduct a power calculation by constructing uniformly distributed observations among the budget lines given by the CentERpanel dataset for both scenarios – taking into account the full set of observations and when we exclude the worst observation (we do not construct uniformly distributed observations for the scenario when the two worst observations are excluded because of computational intensity). Consequently, we compare uniformly generated observations with actual observations. As a result, we show the optimal (in terms of Beatty & Crawford (2011) approach described earlier) fraction of subjects that satisfies the optimal threshold of the CCEI Index for both scenarios – taking into account the full set of observations and when the worst observation is excluded (see section 4 figures 7, 8, and 9 for the results).

4 Combining the Critical Cost Efficiency and Hautman Maks Indices

In this section, we provide various results indicating the positive impact of removing the worst observation per subject on the efficiency of the data set. Furthermore, we provide figures showing the relationship between the CCEI and HM indices. Figure 3 displays the relationship between the mean of the CCEI Index of subjects with respective HM Index. The solid line shows the mean when the full set of observations is present. The vertical dashed line indicates the average HM Index of the sample when the full set of observations is included in the analysis.

Figure 3: CCEI Index Depending on Number of Removed Choices



Source: Author's computation based on CentERpanel dataset.

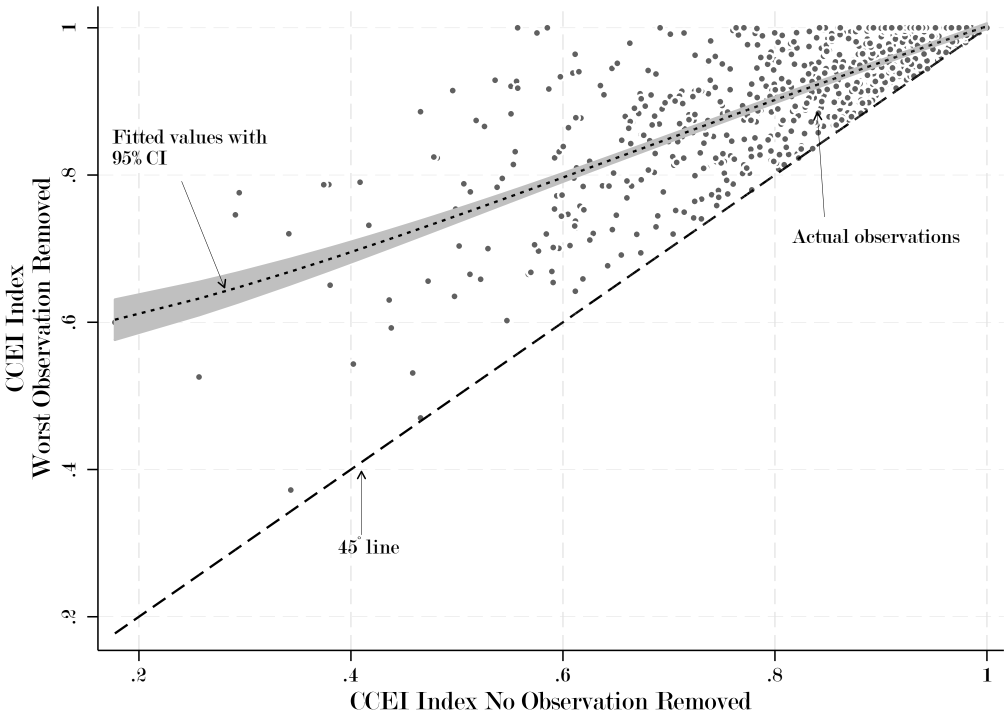
In our case, on average, subjects indicate almost 3 inconsistent choices with respect to GARP. Furthermore, the average CCEI Index (horizontal dashed line) of the sample is 0.88. Put differently, on average, subjects would have to lower their budget by around 12% to be fully consistent with GARP. Moreover, figure 3 shows the mean of the CCEI Index when the worst observation per each subject is removed (dashed line). The shape of the dashed line indicates that we could increase the average CCEI index by almost 6 pp by simply allowing the subject to make one significant mistake. In other words, to achieve the higher threshold of the CCEI Index (in our case, 0.94, represented by the horizontal dashed line), we can exclude the worst observation per subject instead of excluding 'inconsistent subjects' from the sample. Therefore, further study could use observations from more subjects, which are more powerful in terms of the CCEI Index.[9]

[9]Figure 7 shows fraction of subjects depending on the CCEI Index level, however, we describe the figure 7 in more details further down.

Finally, the dotted line indicates the relationship between the mean of the CCEI Index with respective HM Index when the two worst observations per subject are excluded. In this case, the increase in the CCEI Index is even higher than excluding only the worst observation. Nevertheless, the figure shows that the distance between the dashed line and the dotted line is lower than the distance between the solid line and the dashed line. This indicates that removing the worst observation per subject is a suitable approach to achieve an 'efficient' dataset.

The next figure (figure 4) shows the actual increase in the CCEI Index per subject when the worst observation per subject is excluded. The 45° degree line indicates equality between the CCEI indices when no observation and the worst observation are removed. Naturally, by excluding the worst observation, the CCEI Index can only be equal or higher to the CCEI Index when leaving the full set of observations. Therefore, all observations appear above the 45° line.

Figure 4: Scatter Plot Using Fraction Polynomials Fitted Values



Source: Author's computation based on CentERpanel dataset.

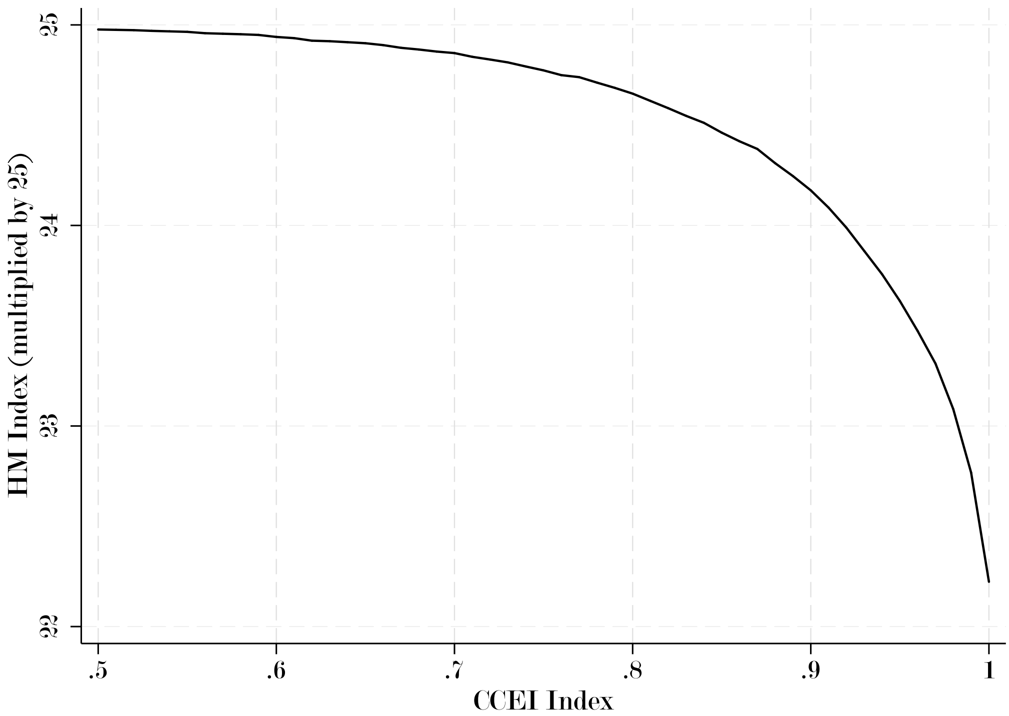
The dotted line indicates the fitted values using the fraction polynomials regression method (with degree 2) and its 95% confidence interval. Figure 4 unambiguously displays the remarkable increase in the CCEI Index when the worst observation per subject is excluded. Moreover, the dotted line shows that the marginal increase in the CCEI Index (when the worst observation is removed) is decreasing with a higher value of the CCEI Index (when no observation is removed).[10]

The further relationship between HM Index and CCEI Index is illustrated in figure 5. The figure displays an intuitive relationship between the HM Index and the CCEI Index – to obtain a higher CCEI index, one must exclude more choices violating GARP.

[10]The decreasing manner is captured in figure 10 in Appendix.

Put differently, given the current dataset, on average, subjects would have to 'throw away' almost 3 observations in order to use their whole budget efficiently. Furthermore, figure 5 suggests that, on average, to use all 25 observations and satisfy GARP, one would have to lower the budget by almost 50 percent. Therefore, the figure shows the exact HM Index when one would like to achieve the minimum CCEI Index level stated on the horizontal line.

Figure 5: Relationship between HM Index and CCEI Index



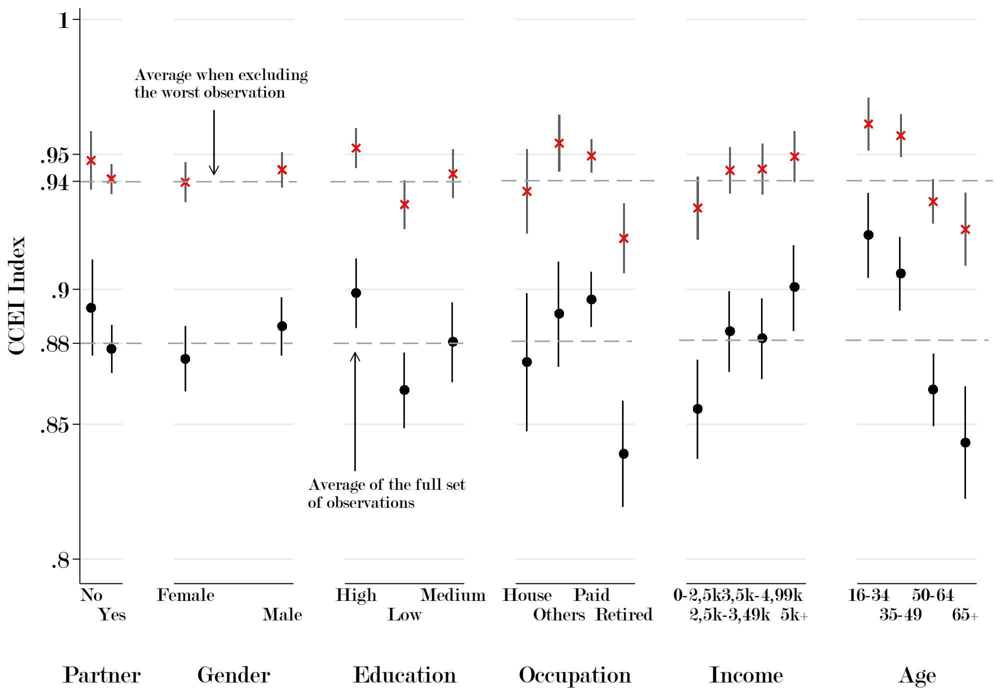
Source: Author's computation based on CentERpanel dataset.

As a next step, we apply the aforementioned findings on the socio–economic groups in the dataset. Figure 6 demonstrates the average increase in the CCEI Index among the socio–economic groups when the worst observation is excluded (x symbols) in comparison with the CCEI Index of the full set of observation (black dots – the results for the full set of observations are replication of the study done by Choi *et al.* (2014)). Taking the full set into account (black dots), on average, high–income and high–educated subjects perform better than lower–educated and lower–income subjects. For example, subjects with an income (in Euro currency) of over 5k+ 'waste' about 10% of their budget compared to almost 14% for those with an income of 0–2.5k. Furthermore, younger subjects display a higher consistency than older subjects, and men tend to maximize their utility more than women. Retired subjects indicate the lowest consistency with respect to GARP. Finally, subjects living with a partner display a lower consistency then those living without a partner.[11] When we exclude the worst observation per each subject, a similar pattern per each group holds (except in the 'Occupation' block, where the group 'Others' obtains the highest value; and in the 'Income' block, where subjects with income of 3.5k–4.99k are now in the second position regarding the CCEI

[11]Figure 11 in Appendix shows the HM Index across socio–economic groups. The pattern is very similar to the CCEI Index.

Index). However, the distance between the highest and the lowest CCEI Index in each block decreases, and, on average, the CCEI Index is higher by 6 pp. As a consequence, the variance of observations decreases and the results have smaller confidence intervals. However, the distance between the observations in groups decreases, and therefore, the statistical significance is almost the same.

Figure 6: CCEI Index across Socio-economic Groups with 95% Confidence Intervals



Source: Author's computation based on CentERpanel dataset.

Table 2 summarizes figure 1 by showing the growth of the CCEI Index among the various socio–economic groups when we exclude the worst choice per subject. Retired subjects indicate the highest growth in CCEI Index, namely 8 pp. This group is followed by the older subjects, aged 65+, with the second highest growth in CCEI Index, namely 7.9 pp. On the other hand, younger subjects, age 16-34, display the lowest increase in CCEI Index, only 4.1 pp. This suggests that, on average, the worst choices made by older subjects are worse than the worst choices made by younger subjects. Other groups indicate similar growth in CCEI Index, around 6 pp.

Table 2: CCEI Index Growth among Various Socio–economic Groups

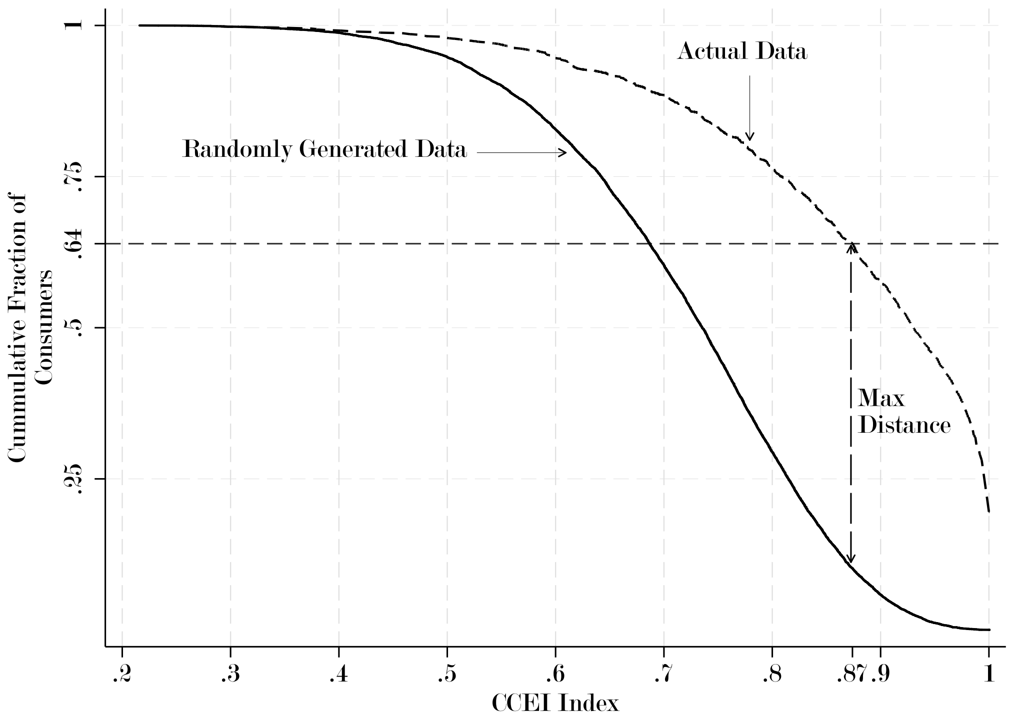
	Percentage Point Increase
Gender	
Female	5.8
Male	6.5
Age	
16-34	4.1
35-49	5.1
50-64	7.0
65+	7.9
Education	
High	5.4
Medium	6.2
Low	6.9
Income	
5k+	4.8
2,5k-3,49k	6.0
3,5k-4,99k	6.3
0-2,5k	7.4
Occupation	
Paid	5.3
Others	6.3
House	6.3
Retired	8.0
Partner	
No	5.5
Yes	6.3

Source: Author's computation based on CentERpanel dataset.

Following the Bronars (1987) and Beatty & Crawford (2011) approach, we conduct a power calculation by constructing uniformly distributed choices using the budget lines given by the CentERpanel dataset. Figures 7, 8, and 9 show the fraction of subjects with respective CCEI Index for the full set of observations, the CCEI Index when we remove the worst observation per subject, and the HM Index, respectively. Starting with figure 7, the solid line shows randomly generated choices among the given budget lines. For instance, almost 10% of random choice sets achieve a CCEI Index of 0.87 (optimal value of CCEI Index based on Beatty & Crawford (2011) approach). The curved dashed line indicates the actual choices from the dataset, almost 64% of subjects achieve the optimal CCEI Index of 0.87. The power of the data (denoted as $\hat{m}_{CCEI}$ in section 3) is illustrated by 'Max Distance' dashed line and it equals to 0.54. Put differently, the actual data outperforms the randomly generated data by 54 pp in terms of fraction of subjects satisfying the optimal CCEI Index. Moreover, the curved dashed line appears above the solid line for each level of CCEI Index. Therefore, the actual choices outperform the randomly generated data by a not–so negligible fraction. Figure 8 shows the same calculation as figure 7, although in this case we exclude the worst observation per subject. In general, the graph in figure 8 'shifts to the right' in comparison to the graph in figure 7. This suggests an overall increase in the efficiency of the data. For instance, taking the randomly generated data into account (solid line), almost 12% of random choice sets achieve a CCEI Index of 0.936 (remarking the optimal level of CCEI Index based on the Beatty & Crawford (2011) approach) compared to 68% (horizontal dashed line) of the actual choices indicated by curved dashed line. In this case, the power of the data (illustrated by 'Max Distance' dashed line) equals to 0.56 (actual choices outperform the randomly generated data by 56 pp regarding the optimal CCEI Index level). Therefore, by excluding the worst observation per subject we achieve higher 'optimum' CCEI Index with larger fraction of subjects having higher power ($\hat{m}_{CCEI}$) than leaving the full set of observations.

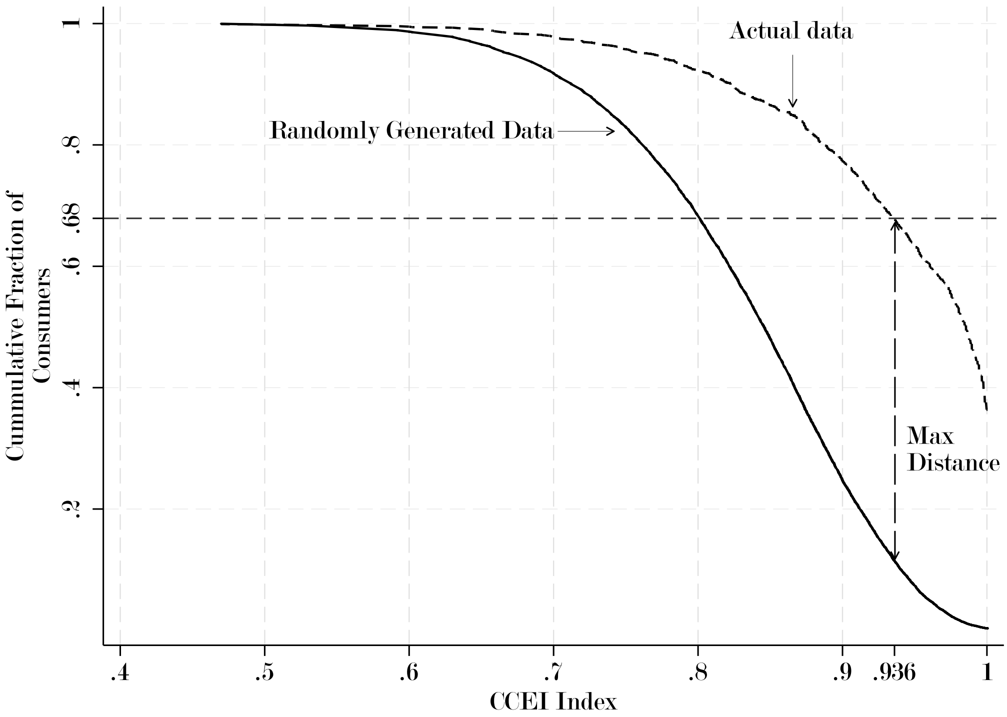
Figure 9 illustrates a similar approach as the one used in figure 7 (using the full set of observations), but with the HM Index on the x axis. Using the Beatty & Crawford (2011) approach, we calculate the 'optimum' HM Index with respective fraction of subjects. As a result, almost 10% of randomly generated choices (solid line) achieve the 'optimum' HM Index equal to 22. On the other hand, 64% of subjects with actual choices (horizontal dashed line) achieve an HM Index of 22. The power of the data (indicated by the 'Max Distance' dashed line) is equal to 0.54 (meaning that the actual data outperforms the randomly generated data by 54 pp regarding the optimal level of HM Index). Therefore, the result suggests that the actual choices strongly outperform randomly generated data.

Figure 7: CCEI Index of Randomly Uniformly Distributed Data vs Actual Data



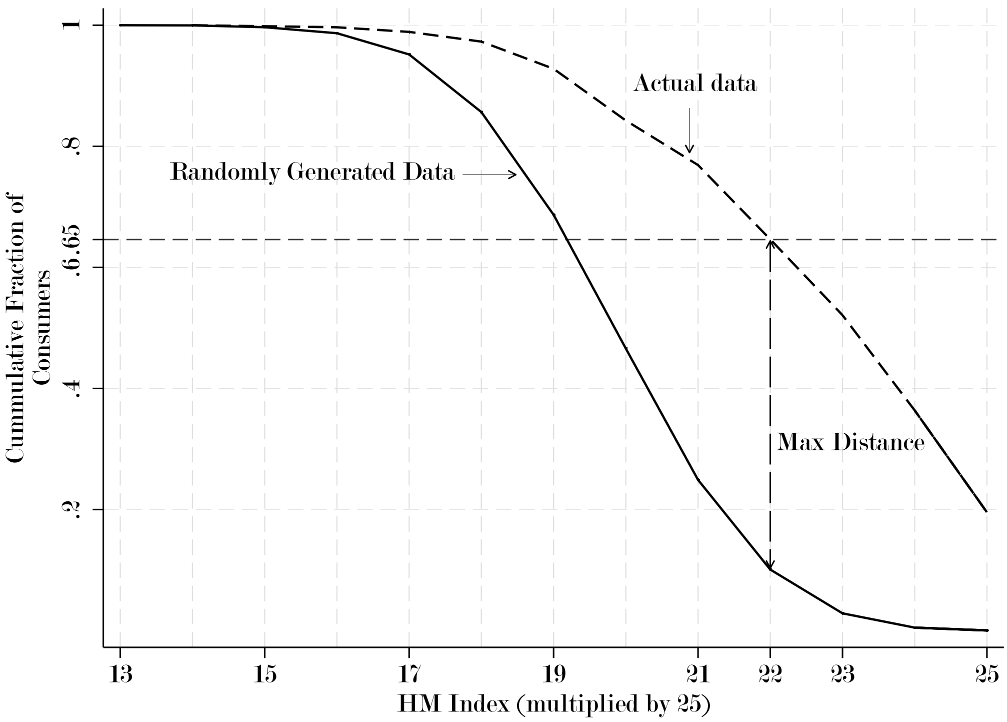
Source: Author's computation based on CentERpanel dataset.

Figure 8: CCEI Index of Randomly Uniformly Distributed Data vs Actual Data – Removing the Worst Observation



Source: Author's computation based on CentERpanel dataset.

Figure 9: HM Index of Randomly Uniformly Distributed Data vs Actual Data



Source: Author's computation.

Finally, table 3 summarizes the fraction of subjects satisfying a certain level of CCEI Index when both scenarios are present; the full set of observations and when we exclude the worst observation per subject. The results indicate that the fraction of subjects increase by 19 pp on the CCEI Index level of 0.95, 0.90, 0.85. Moreover, CCEI Index levels of 1.00 and 0.80 indicate a slightly lower, but notable increase in the fraction of subjects (+16 pp).

Table 3: Fraction of Subjects per CCEI Index

CCEI Index	Full set of observations	The worst observation excluded
1.00	0.20	0.36
0.95	0.45	0.64
0.90	0.58	0.77
0.85	0.68	0.87
0.80	0.76	0.92

Source: Author's computation based on CentERpanel dataset.

Therefore, taking into account the evidence from figures 3, 4, 6, 7, 8 and 9, excluding the worst observation per subject in the dataset indicates an efficient way to increase the fraction of subjects that satisfy a certain CCEI Index threshold.

5 Conclusion

In this thesis, we use the CentREpanel dataset to examine the relationship between CCEI and HM indices. Using the algorithm developed by Gross & Kaiser (1996) and Houtman & Maks (1985), we calculate the CCEI and HM indices among various socio–economic groups. Furthermore, we identify the worst observation per subject and we demonstrate the effect of excluding the worst observation per subject on the level of CCEI Index.

Our analysis shows that to obtain a CCEI Index equal to one (meaning that each subject would use his/her budget efficiently), on average, subjects would have to 'throw away' almost 3 observations. Moreover, this study demonstrates an average increase of the CCEI Index of 6 pp when we exclude the worst observation per subject in our sample. Additionally, we determine the effect of removing the worst observation among various socio economic groups. The most notable increase in the CCEI Index is among the 65+ aged, and retired subjects. On the other hand, the youngest subjects indicate the lowest increase of the CCEI Index. In addition, we employ Bronars (1987) and Beatty & Crawford (2011) approaches to calculate the optimal CCEI Index level with respective fraction of subjects satisfying the optimal level of CCEI Index. We demonstrate that by excluding the worst choice per subject not only does the optimal CCEI Index level increase, but a larger fraction of subjects satisfies this optimal CCEI Index level. Furthermore, we present a remarkable increase in the fraction of subjects satisfying various levels of CCEI Index (we choose levels of 1.00, 0.95, 0.90, 0.85, and 0.80). In all cases, the increase is between 16 pp to 19 pp. This suggests that by excluding the worst observation per subject one would increase the 'efficient' fraction of the dataset by an important amount. Therefore, one could test the theory with a more consistent dataset. Additionally, more subjects would pass the threshold of CCEI Index and, as a result, one would have to 'throw away' less data.

Overall, this study shows that combining CCEI and HM indices, and as a consequence, excluding the worst observation per subject, indicates an efficient way to increase the consistency of the data. However, the aforementioned findings are based on one dataset. Therefore, further study needs to be done to fully address the efficiency of data in general.

References

Afriat, S. N. (1967): "The construction of utility functions from expenditure data." *International economic review* **8(1)**: pp. 67–77.

Afriat, S. N. (1972): "Efficiency estimation of production functions." *International Economic Review* pp. 568–598.

Ahn, D., S. Choi, D. Gale, & S. Kariv (2014): "Estimating ambiguity aversion in a portfolio choice experiment." *Quantitative Economics* **5(2)**: pp. 195–223.

Andreoni, J. & J. Miller (2002): "Giving according to garp: An experimental test of the consistency of preferences for altruism." *Econometrica* **70(2)**: pp. 737–753.

Banerjee, S. & J. H. Murphy (2006): "A simplified test for preference rationality of two-commodity choice." *Experimental Economics* **9(1)**: pp. 67–75.

Banerjee, S. & J. H. Murphy (2011): "Do rational demand functions differ from irrational ones? evidence from an induced budget experiment." *Applied Economics* **43(26)**: pp. 3863–3882.

Battalio, R. C., J. H. Kagel, R. C. Winkler, E. B. Fisher, R. L. Basmann, & L. Krasner (1973): "A test of consumer demand theory using observations of individual consumer purchases." *Economic Inquiry* **11(4)**: p. 411.

Beatty, T. K. & I. A. Crawford (2011): "How demanding is the revealed preference approach to demand?" *The American Economic Review* **101(6)**: pp. 2782–2795.

Becker, N., K. Häger, & J. Heufer (2013): "Revealed notions of distributive justice ii: Experimental analysis." *Ruhr Economic Papers* **444**.

Bronars, S. G. (1987): "The power of nonparametric tests of preference maximization." *Econometrica: Journal of the Econometric Society* pp. 693–698.

van Bruggen, P. & J. Heufer (2017): "Afriat in the lab." *Journal of Economic Theory* **169**: pp. 546–550.

Bruyneel, S., L. Cherchye, S. Cosaert, B. De Rock, & S. Dewitte (2012): "Are the smart kids more rational?" .

Burden, R. L. & J. D. Faires (1985): "2.1 the bisection algorithm." *Numerical Analysis. Prindle, Weber & Schmidt, Boston, MA., pp. x* **676**.

Burghart, D. R., P. W. Glimcher, & S. C. Lazzaro (2013): "An expected utility maximizer walks into a bar..." *Journal of risk and uncertainty* **46(3)**: pp. 215–246.

Camille, N., C. A. Griffiths, K. Vo, L. K. Fellows, & J. W. Kable (2011): "Ventromedial frontal lobe damage disrupts value maximization in humans." *Journal of Neuroscience* **31(20)**: pp. 7527–7532.

Choi, S., R. Fisman, D. Gale, & S. Kariv (2007): "Consistency and heterogeneity of individual behavior under uncertainty." *The American economic review* **97(5)**: pp. 1921–1938.

Choi, S., S. Kariv, W. Müller, & D. Silverman (2014): "Who is (more) rational?" *The American Economic Review* **104(6)**: pp. 1518–1550.

Dawes, C. T., P. J. Loewen, & J. H. Fowler (2011): "Social preferences and political participation." *The Journal of Politics* **73(3)**: pp. 845–856.

Dickinson, D. L. (2009): "Experiment timing and preferences for fairness." *The Journal of Socio-Economics* **38(1)**: pp. 89–95.

Février, P. & M. Visser (2004): "A study of consumer behavior using laboratory data." *Experimental economics* **7(1)**: pp. 93–114.

Fisman, R., S. Kariv, & D. Markovits (2007): "Individual preferences for giving." *The American Economic Review* **97(5)**: pp. 1858–1876.

Gross, J. & D. Kaiser (1996): "Two simple algorithms for generating a subset of data consistent with warp and other binary relations." *Journal of Business & Economic Statistics* **14(2)**: pp. 251–255.

Harbaugh, W. T. & K. Krause (2000): "Children's altruism in public good and dictator experiments." *Economic Inquiry* **38(1)**: pp. 95–109.

Harbaugh, W. T., K. Krause, & T. R. Berry (2001): "Garp for kids: On the development of rational choice behavior." *The American economic review* **91(5)**: pp. 1539–1545.

Heufer, J. & P. Hjertstrand (2015): "Consistent subsets: Computationally feasible methods to compute the houtman–maks-index." *Economics Letters* **128**: pp. 87–89.

Houtman, M. & J. Maks (1985): "Determining all maximal data subsets consistent with revealed preference." *Kwantitatieve methoden* **19**: pp. 89–104.

Mattei, A. (2000): "Full-scale real tests of consumer behavior using experimental data." *Journal of Economic Behavior & Organization* **43(4)**: pp. 487–497.

Selten, R. (1991): "Properties of a measure of predictive success." *Mathematical Social Sciences* **21(2)**: pp. 153–167.

Sippel, R. (1997): "An experiment on the pure theory of consumer's behaviour." *The Economic Journal* **107(444)**: pp. 1431–1444.

Tversky, A. & D. Kahneman (1975): "Judgment under uncertainty: Heuristics and biases." In "Utility, probability, and human decision making," pp. 141–162. Springer.

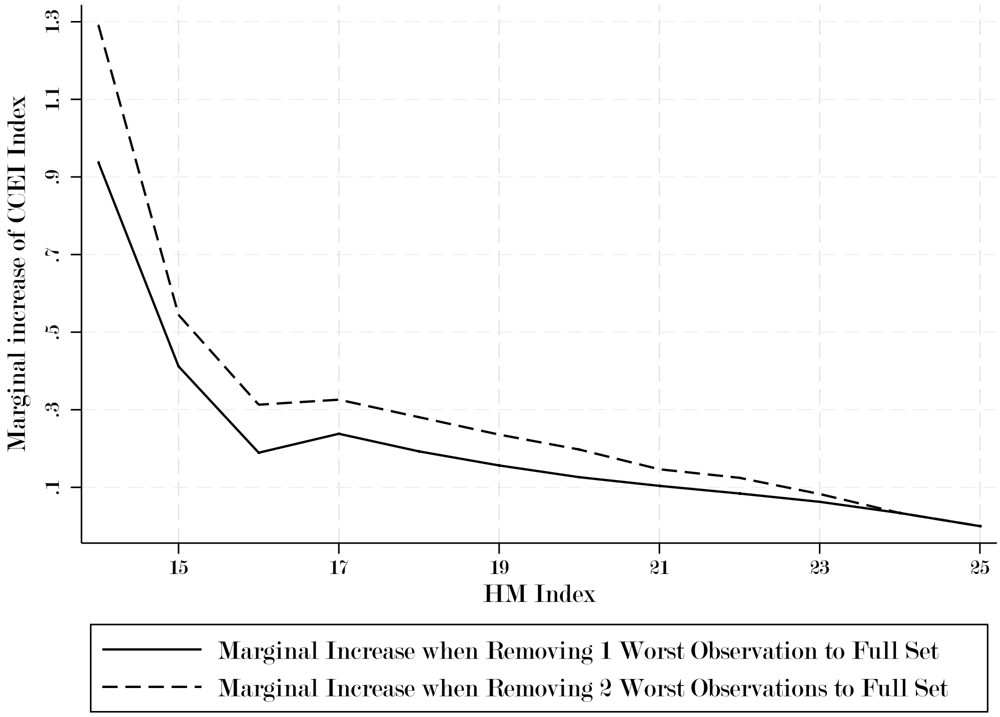
Varian, H. R. (1982): "The nonparametric approach to demand analysis." *Econometrica: Journal of the Econometric Society* pp. 945–973.

Varian, H. R. (1991): *Goodness-of-fit for revealed preference tests.* Department of Economics, University of Michigan.

Visser, M. S. & M. R. Roelofs (2011): "Heterogeneous preferences for altruism: gender and personality, social status, giving and taking." *Experimental Economics* **14(4)**: pp. 490–506.

Appendix

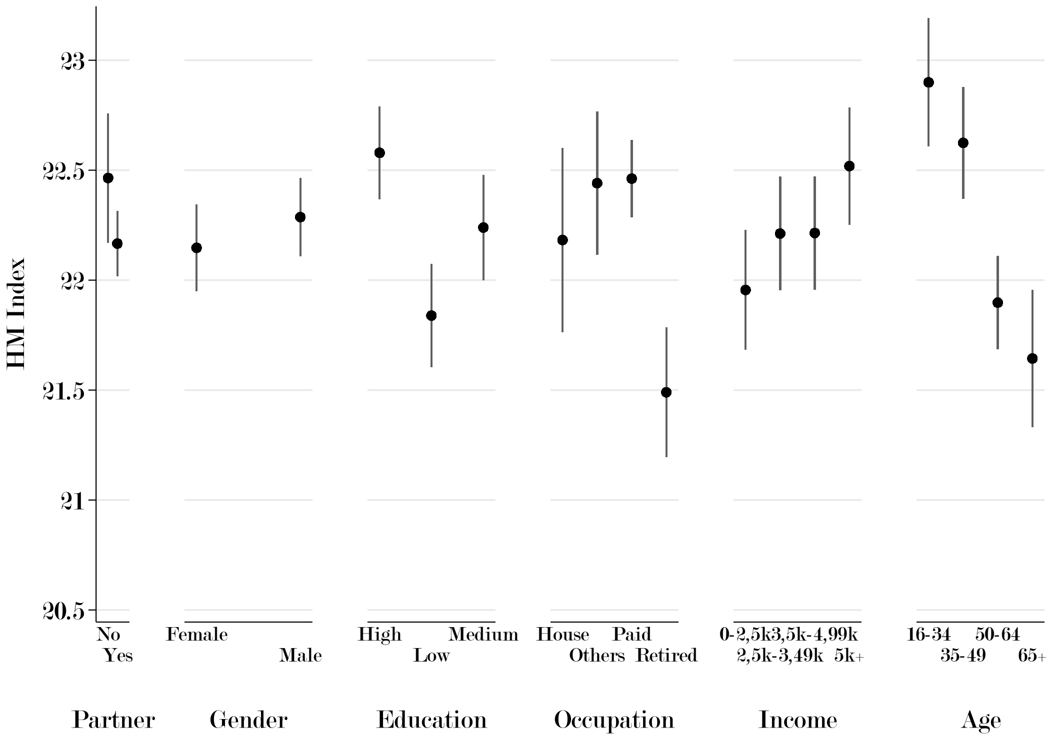
Figure 10: Marginal Increase in CCEI Index when Removing One or Two Choices



The figure indicates that the marginal increase in CCEI Index decreases with higher level of HM Index.

Source: Author's computation based on CentERpanel dataset.

Figure 11: HM Index across Socio-economic Groups



Source: Author's computation based on CentERpanel dataset.

The Matlab codes, Stata codes for graphs, and dataset can be found on:

https://drive.google.com/open?id=0BxOyI0hgC0n2OXI4ZlFWSlFXcDA