

**CHARLES UNIVERSITY IN PRAGUE**

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

**Bachelor thesis**

**2012**

**Jana Malíčková**

CHARLES UNIVERSITY IN PRAGUE

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies



Jana Malíčková

Archetypal analysis of students' behavior  
in banking sector

*Bachelor thesis*

Prague 2012

Author: **Jana Malíčková**

Supervisor: **RNDr. Michal Červinka, Ph.D.**

Advisor: **RNDr. Pavel Ranocha, Ph.D.**

## **Bibliografický záznam**

MALÍČKOVÁ, Jana. *Archetypal analysis of students' behavior in banking sector* Praha 2012. 50 s. Bakalářská práce, Univerzita Karlova, Fakulta sociálních věd, Institut ekonomických studií. Vedoucí bakalářské práce: RNDr. Michal Červinka, Ph.D. Konzultant bakalářské práce: RNDr. Pavel Ranocha, Ph.D.

## **Abstrakt**

Archetypální analýza je metoda používaná pro analýzu segmentace trhu, avšak nebyla doposud mnohokrát použita v ekonomii. Tato metoda se snaží nalézt typická chování – archetypy. Každý jedinec se skládá z jednoho či více archetypů a každý archetyp se skládá z několika jedinců. V práci zkoumáme chování studentů českých vysokých škol v bankovním sektoru, jelikož doposud nebyla žádná zmínka o takovém průzkumu. V modelu se čtyřmi archetypy vznikly archetypy pojmenované Chud'as, Buffett, Elvis a Šípková Růženka.

**Klíčová slova:** archetypální analýza, bankovníctví, anketní šetření

**E-mail autora:** janeet@seznam.cz

**E-mail vedoucího práce:** cervinka@utia.cas.cz,

**E-mail konzultanta práce:** pavel.ranocha@tnsglobal.com

## **Abstract**

Archetypal analysis is statistical method used in market segmentation, however, it has minor results of application in economics. This method finds out typical behavior – archetypes. Every sample point consists of one or more archetypes and every archetype consists of few sample points. In this thesis, we explore behavior of college students at Czech universities in banking sector as there is no evidence of this research. In the model with four archetypes, the archetypes named the Pauper, the Buffett, the Elvis and the Sleeping Beauty arose.

**Keywords:** archetypal analysis, banking, survey

**Author's e-mail:** janeet@seznam.cz

**Supervisor's e-mail:** cervinka@utia.cas.cz,

**Advisor's e-mail:** pavel.ranocha@tnsglobal.com

## **Declaration of Authorship**

I hereby proclaim that I wrote my bachelor thesis on my own under the leadership of my supervisor and advisor and that the references include all resources and literature I have used.

I further declare that the thesis has not been used previously for obtaining any university degree.

I grant a permission to reproduce and distribute copies of this thesis document in whole or in part. I do not grant a permission reproduce or distribute copies of the numerical software as I do not have property rights.

## **Acknowledgment**

I would like to express my gratitude to RNDr. Michal Červinka, Ph.D. for his patience, time, valuable notes and guidance.

I would also like to thank to RNDr. Pavel Ranocha, Ph.D. for providing me with his numerical software and for his valuable notes.

I am also grateful to respondents of my survey.

Prague, May 16, 2012

---

Signature

# Bachelor Thesis Proposal

The aim of the thesis is to explore students' behavior in banking sector using archetypal analysis.

Archetypal analysis represents sample points through a combination of few archetypes, i.e. individuals with representative attributes. An integral part of analysis is not only identification of archetypes in a given data set and their interpretation, but also creation of survey and data collection. One of the main current applications of archetypal analysis applied on demand and consumption of a given product is the improvement of marketing strategy of producers and creation of advertisement aimed at the representative customers.

The author will explore students' behavior in banking sector. Using survey with a sufficient number of college students, she will get information about how and what banking products students use. In a suitable software, the data will be processed and using archetypal analysis will be identified archetypes and their deep description.

The author will consult the thesis with external supervisor RNDr. Pavel Ranocha, Ph.D., from the TNS AISA company.

## **Preliminary structure of the thesis:**

1. Brief description of archetypal analysis
2. Description of survey of students' behavior in banking sector
3. Overview of used algorithm in suitable software
4. Analysis of the results and interpretation

## **Literature:**

1. A. Cutler, L. Breiman: Archetypal Analysis, *Technometrics*, 36(4), 338-347, 1994.
2. M.J.A. Eugster, F. Leisch: From Spider-Man to Hero Archetypal Analysis in R, *Journal of Statistical Software*, Volume 30, Issue 8, 2009. (available online at <http://www.jstatsoft.org/v30/i08/paper>)

# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction</b>  | <b>1</b>  |
| <b>2</b> | <b>Market targeting methods</b>                                      | <b>4</b>  |
| <b>3</b> | <b>Brief description of archetypal analysis</b>                      | <b>7</b>  |
| 3.1      | Examples . . . . .   | 9         |
| 3.2      | Mathematical background . . . . .                                    | 11        |
| <b>4</b> | <b>Overview of used algorithm in suitable software</b>               | <b>16</b> |
| <b>5</b> | <b>Description of survey of students' behavior in banking sector</b> | <b>18</b> |
| <b>6</b> | <b>Analysis of the results and their interpretation</b>              | <b>25</b> |
| 6.1      | Model with 1 archetype – Average Respondent . . . . .                | 25        |
| 6.2      | Model with 2 archetypes . . . . .                                    | 26        |
| 6.3      | Model with 4 archetypes . . . . .                                    | 27        |
| <b>7</b> | <b>Summary</b>   | <b>32</b> |
|          | <b>References</b>  | <b>33</b> |
| <b>A</b> | <b>Appendix</b>  | <b>37</b> |
| A.1      | Results of the survey and recoding . . . . .                         | 37        |
| A.2      | SSE . . . . .  | 41        |
| A.3      | Model with 1 archetype . . . . .                                     | 42        |
| A.4      | Model with 2 archetypes . . . . .                                    | 43        |
| A.5      | Model with 4 archetypes . . . . .                                    | 45        |

# 1 Introduction

*“Young people think money is important. Old people know it.”*

doc. Ing. Karel Půlpán, CSc.

The importance of market research has grown up in the past 100 years as the competitiveness among producers increased. The aim of market research is to analyze potential consumers, their needs and behavior, choose the most suitable one for a market strategy of a given company. There are a lot of techniques included in market research. Many psychological, advertising, sociological and neuroscientific tools are used in this branch of economics. Also mathematics and its tools should not be disregarded. In market research, many statistical and econometric procedures are used, such as cluster analysis, data mining, regression etc.

In 1994, statistical archetypal analysis was introduced by Adele Cutler and Leo Breiman in [3]. At first, archetypal analysis was used in natural sciences like biology and physics. Later, this method was also applied in social sciences which usually work in terms of a qualitative analysis. This approach has recently found a utilization also in economics. We would like to explore this area in this thesis.

There is no evidence of a research of students' behavior in banking sector in the Czech Republic. Nevertheless, few surveys about financial behavior of a whole population have been conducted in the world as well as in the Czech Republic. Authors of the surveys are mostly public sector (government of academic researches), banking institutions or non-profit organizations. The surveys are connected with financial literacy, education and behavior (planning family budget, retirement saving and consumer protection in signing contracts).

ING Bank conducted an online survey among 5,000 respondents across 10 countries in Europe, the U.S., Latin America and Asia in 2010. The Czech Republic was not included. The main topics consisted of financial literacy, attitudes, motivations and money skills.

According to [12], young people are significantly more financially literate than older people and men are more financially literate than women. The research also shows that Japanese, Indians and South-Koreans have quite high financial literacy, on the other hand, there is a low level of financial literacy in Mexico, Romania and Spain. It was also explored that 65% of respondents had emergency fund that would not cover their expenditures for more than 3 months. The survey does not explore usage of financial products.

In the Czech Republic, the Ministry of Finance and the Czech National Bank in collaboration with the agency STEM/MARK conducted a face to face survey about the level of financial literacy in the Czech population in 2007 and also in 2010. The survey from 2010 included a wide range of respondents – 1005. In terms of age, there were included 20% of people younger 29 years. The survey shows that people know a lot of financial products, but they use only few of them – 68% have a bank account (66% of people under 29 years old have a bank account), 48% have a pension insurance (according to [15], there are 4,565,741 contracts as of 2011), 37% have building savings (according to [16], there are 4,550,468 contracts as of 2011), 35% have a debit card and 24% have a credit card. 33% of respondents use Internet Banking. It is no surprise that people with higher education are more financial responsible and literate, but the general level of financial literacy is low. As a result, the Ministry of Finance created a Financial Education Strategy.

In terms of financial behavior of students, surveys about students and loan have been recently conducted. According to [10], since late 1990s the concern about the rising use of credit cards among college students has gone up as in the U.S. college students pay tuition and fees. According to [2], college students have on average two credit cards and their debt level is between USD 1,000 and 2,000. In the Czech Republic, the college students in public colleges do not have to pay tuition (the situation in 2012). Hence, in our country college students will most likely have a lower level of debt and also other behavior in banking sector.

Banking products for students, such as bank accounts, credit and debit cards, are not differentiated in the Czech Republic, even though we believe that behavior of students of Czech college differs – some prefer consumption, some else savings, some have high income, some else low income, some have to pay necessary expenditures, etc. Despite the fact that students are important customers for banks as they will demand mortgages, loans, which generate profits for banks, in short term, banks usually offer only one standardized type of a bank account for students. If banking institution focuses on one or few specific segments of students, it could generate higher profits.

In our survey, we would like to focus on whether, how and to what extent Czech college students use banking and financial products. Do they prefer consumption or savings? What is their attitude to money, wealth and investments? Using archetypal analysis we are going to find out some pure types – archetypes – to confirm the hypothesis that students do not behave in the same way, but there can be found many differences in college students' behavior. And also find out whether young people think that money is important.

This bachelor thesis is organized as follows. First, we are going to present market targeting methods which are commonly used. In Section 3, we briefly describe archetypal analysis – what archetypes mean, where we can meet archetypes, examples of archetypal analysis and mathematical background of this method. Overview of used software will be presented in Section 4. Then we describe our survey of students' behavior in banking sector. In Section 6, we will analyze results of archetypal analysis, present and describe arisen archetypes. Appendix contains detailed results of our survey and detailed descriptions of archetypes.

## **2 Market targeting methods**

Market consists of many consumers, producers, retailers, institutions, goods, services and their interactions. Consumers differ from each other as they have different age, gender, income, needs, behavior and attitudes. Furthermore, producers can have different goals (profit maximization, business survival, goodwill, etc.), although, profit maximization predominates. Microeconomics studies the behavior of an individual or a company, how to allocate limited resources and maximize profit. Market research deals with reaching goals of a given producer and determining which market segment offers the best opportunity. Many market research methods to target costumers have been already considered. In this section, we briefly introduce just a few of them.

### **Mass marketing**

Mass marketing is the first type of a coverage strategy in the history. It is also called undifferentiated marketing and has its origins in the 1920's. This marketing method "One Size Fits All" ignores differences between consumers. It uses same distribution methods, same channels, same advertising (often radio, television and newspaper) and only one product. Even though it creates the largest potential market, it is hard to produce one product, which fits all consumers. Products like durable goods (furniture, cars, personal computers, washing machines, etc.) are subject to mass marketing.

### **Segmentation marketing**

Market segmentation means dividing overall market into two or more homogeneous groups that distinguishes customers' needs, characteristics and behavior. Each group has a different marketing mix, according to its gender, age, income, location, etc. A fine tuning of products works more efficiently than mass marketing as consumers respond more readily if offered products are close to their needs.

There are many advantages of this method. It can help in developing a new product, in creating a distributional pricing and promotional tactics. This method also enables producers to react on changes in demand more elastically and it also reduces the possibility of a competition. Finally, the market segmentation leads to more efficient cost allocation of a given producer.

According to [11], the market segmentation is conducted in three phases. First, a research about customers is carried out. It aims to collect data about desired features of a product, characteristics of customers, attitudes to a kind of goods and services. Then an analytical phase follows, where the collected data are statistically analyzed. The third phase, profiling, consists of characterization of segments, which have arisen in the analytical stage.

### **Personalized marketing**

Personalized marketing can be considered as an opposite marketing strategy of mass marketing. Its basic idea consists of a belief that each consumer is different and makes a unique product for each customer. Hence, it is sometimes called one-to-one marketing. This method is practical with the internet as this interactive media can track interests of each user, especially with Facebook, where we can give “likes” to pages, status and photos or comment them. One-to-one marketing does not increase the number of customers, but it increases the value of a customer base.

According to [17], there are four stages in putting one-to-one marketing to work. First, one has to identify customers of a given company and try to get information about them (names, addresses, habits, preferences, etc.). A second stage is called differentiating customers, which consists of selecting the most valuable customers. Then company’s behavior to each customer could be tailored in order to reflect customer’s needs. Third, interacting with customers improves cost-efficiency by more automated and cheaper channels. Also effectiveness is improved as this method provides better insight into customer. Fourth, customizing enterprise’s behavior – in this stage a company

must adapt some aspects of its behavior to satisfy customer's individually expressed needs.

### **Niche marketing**

Niche marketing focuses on segments which are new and where is almost no competition. The word *niche* comes from the French word that means "to nest". Its aim is to fill up a space in the market. Benefits of niche marketing consists of a unique product that other may not offer, a unique brand which appeals to a particular target audience, and continuous income – since producer will supply to a small group of customers, producer will likely get a certain amount, but since competition is not tight, it will ensure a continuous income.

There are many other market research methods: relationship marketing, Internet marketing, real-time marketing, etc.

The previous examples represent approaches of companies to customers. There are several methods how to conduct market analysis. One of these methods is called archetypal analysis.

### **Archetypal analysis**

Archetypal analysis can be used for mass marketing, segmentation marketing and personalized marketing as well. As we would like to find some archetypes, we are interested in segmentation. However, archetypal analysis does not mean dividing into homogeneous groups – in this specific method, we try to find out typical pure, or maybe we could say extremal, behavior. Every sample point is a convex linear combination of some or all archetypes. Every archetype is a mixture of few sample points. The advantage of archetypal analysis consists of a very easy description and understanding of segments, the difference between them are evident at first glance. The selection of typical representant for the following analysis is very straightforward. The archetypes are also given accurate names.

### 3 Brief description of archetypal analysis

The word *archetype* is derived from the Latin word *archetypum*, which has origins in the Greek noun *archetypon*. *Arché* means “beginning, origin” and *typos* means “pattern, model, type”<sup>1</sup>. An archetype is defined as the original pattern or model of which all things of the same type are representations or copies<sup>2</sup>.

To easily understand what an archetype is, let us recall fairy tales. There is a limited number of characters in every fairy tale. We associate with each character some attributes even before we know the whole story. The characters tell us more about their holders. Whether they are good or bad, rich or poor, beautiful or ugly, brave or cowards, etc. With a hero we associate bravery, an adventure and an honor (Robin Hood). A stepmother always represents a negative character, she tries to harm her stepdaughter (Cinderella). A princess is often very beautiful and graceful and wants to find her future husband (Beauty and the Beast). A king represents a brightness and an authority. Also with places we associate some stereotypes. Tree and the green color represent knowledge and life. Tunnels, caves and dark places symbolize inactivity, secrets, sometimes death. The sea stands for chaos and danger. Road are compared to journey, adventure or something new.

Not only in the literature for children we can find archetypes, but also in pieces of the world literature appears a limited number of characters, types of relationships and plots. Many characters appear only with some partner (e.g. Romeo and Juliet). Love (mother’s, father’s, partner’s) and fight (for love, for money, for power) in many variations creates a substance of almost all stories.

In the beginning of the 20th century, archetypes were described in the psychology. Carl Gustav Jung, a Swiss psychiatrist and the founder of analytical psychology (or Jungian psychology), focused on dream analysis and symbolism. He described archetypes as innate structural elements of human psyché, inherent possibilities of imag-

---

<sup>1</sup>V. Šolc (2009): *Archetyp otce*, Triton

<sup>2</sup><http://www.merriam-webster.com/dictionary/archetype>

inations, which are common to all of us. Even though archetypes can not be seen, we can see their actions as they produce archetypal images and motives. Archetypal images are themes, a universal patterns of behavior that are derived from a collective or personal unconsciousness. Collective archetypes are basic contents of religion, mythological stories, legends and fairy tales; personal archetypes are patterns of thinking and behavior. Jung framed five main archetypes: the Anima, the Animus, the Persona, the Self and the Shadow.

We do not have to divide individuals only according to their appearance, we can do it also according to their nature and attributes. Greek physician Hippocrates of Cos had already in the ancient Greek developed a theory of body fluids that affect behavior and emotions. Four temperamental categories were named: “choleric”, “melancholic”, “phlegmatic” and “sanguine”. Two aspects (stability/unsteadiness and introverts/extroverts) are considered to determine in which category an individual belongs. A sanguine is a stable extrovert, cheerful, active, sociable and communicative one. A choleric is described as an unstable extrovert, impulsive and restless one. A melancholic is an unstable introvert, unsociable, earnest, pessimistic and rigid one. A phlegmatic is a stable introvert, reliable, quiet, peaceful and indifferent one. An individual is often not a pure type, but a combination of more types, where one type predominates. Using the Eysenck Personality Test we can determine, what type characterizes us at the best way.

We use archetypes and associations on a daily basis. Every day we meet hundreds of people (at least if we live in a big city), which we do not know, even so we have some imaginations and prejudices about them based on our previous knowledge and experience. For example: a man in a suit most likely works in an office in a high position, earns a lot of money and has a fancy car, according to our archetypal prejudice.

We can assign people attributes even based on more circumstantial events such as date of birth. Take a Zodiac as an example, which consists of twelve astrological signs (Aries, Taurus, Gemini, Cancer, Leo,

Virgo, Libra, Scorpio, Sagittarius, Capricorn, Aquarius and Pisces). We assign some typical characteristics to each of the signs, e.g. the Leo (22 July – 22 August) is described as a dominant, extrovert, ambitious and sociable individual. The astrology studies the astrological signs more deeply and builds horoscopes, which are based on the signs. Hence people who believe in astrology plan frequently their action according to their sign.

In 1994, the archetypal analysis was used in mathematics for the first time. The aim of statistical archetypal analysis is to understand a structure of data. Up to now, archetypal analysis was used in physics, the army, biology, etc. (see Section 3.1). There are only minor results of application of archetypal analysis in economics.

### **3.1 Examples**

The following four examples illuminate how broadly archetypes can be used. The variety of areas of application where archetypal analysis could be applied ranges from the army to social sciences through biology.

#### **Biology – Skeletal archetypes**

In [9], the authors took body girth measurements and skeletal diameter measurements and age, height, weight and gender on 247 men and 260 women in their twenties and early thirties, all physically active. They regressed weight on girth measurements and affirmed the hypothesis that body build variables and height predict scale weight substantially better than height alone. In [6], the authors used the data for archetypal analysis of sample points included in the research. The number of calculated archetypes ranged between 1 and 15. In order to find the most suitable number of archetypes, the authors used the “elbow criterion”, which tells us what number of archetypes we should choose such that adding another archetype does not give much better result in a sense explained later in section 3.2. The models with 3 and 7 archetypes were considered as the best ones. In the model with 3

archetypes, archetype 1 represented individuals with average measure except the pelvis measurements. Archetype 2 represented individuals which were “huge” in all measurements. Archetype 3 represented individuals which were “small”.

### **The Army – Swiss Army Head-Dimension Data**

In the mid-80s, 900 members of the Swiss Army were sampled. The data included 200 male soldiers who were twenty years old at the time of investigation. The measurements consisted of the width of the face just above the eyes, the width of the face below the mouth, the distance between the eyes and the chin, the length of the nose, the distance between ear and the top of the nose and one the length between ear and the chin.

The aim of the analysis was to minimize costs of masks production taking into account sizes of soldiers’ heads. In [7], the number of archetypes considered rose from 2 to 5. Having only two archetypes, the “first head” was small and the “second one” was large. Then the rising number of archetypes made the heads more specialized.

### **Representation of religion in countries**

In [5], the author analyzed a representation of religion (Catholic, Islamic, Protestant, etc.) in 72 countries, including the Czech Republic. Cluster and archetypal analysis was used among other types of analysis. In archetypal analysis, there were found 3 and 6 archetypes. In the model with 3 archetypes, Archetype 1 is dominated by Protestant religion (43.8%) in combination with other religions. The typical country with this type of representation of religion is South Africa. In Archetype 2, Catholic religion significantly prevails (92.2%). Portugal is very close to this amount of shares of religion. In Archetype 3, Islamic religion dominates (84.5%) and the typical country is Uzbekistan. In the model with 6 archetypes, there are 4 archetypes that were strongly dominated by one religion (Islamic, Catholic, Old or Protes-

tant). The two other archetypes were dominated by a combination of several religions.

Let us mention that in cluster analysis, the 3 arisen clusters have “more average” share of religions. In this example, we can notice that archetypal analysis does not create segments, but few extremal archetypes (e.g. Portugal, South Africa and Uzbekistan in the model with 3 archetypes). Each sample point (e.g. country) consists of one or more archetypes with a different value.

### **Social sciences – Consumer Assessment of Social Product Features**

In [1], respondents stated whether they would or would not buy shoes described by some attributes. Respondents were City University of Hong Kong undergraduate business students, AGSM MBA students and members of Amnesty International in Australia. Attributes included 8 basic product features (e.g. ankle support, reflectivity at night), 4 ethical features (e.g. child labor in making the product, acceptance of worker’s living conditions at the factory), price and 11 brands of shoes (e.g. Nike, Adidas, Puma, etc.). The authors used 32 responses to calculate archetypes, 4 archetypes were identified as the best number of archetypes. Archetype 1 cared about shoe performance, had a social conscience and favored one specific brand. Archetype 2 focused on price regardless of labor practices. Archetypes 3 was very price sensitive and also concerned about support. Archetype 4 was mainly interested in weight, comfortable fit of shoe, child labor, workers paid minimum wages and one specific brand.

In this example, we can see that archetypal analysis can be used also in social studies where some attributes are not quantitatively expressed.

## **3.2 Mathematical background**

In this section, we briefly describe mathematical background of archetypal analysis, which leads to find out archetypes. We also introduce

a new application of the so-called Herfindahl index for subsequent analysis of archetypes.

For simplicity of notation, every vector is considered as a row vector. For multivariate data  $x_i, i = 1, \dots, n$ , where each  $x_i$  is an  $m$ -dimensional vector  $x_i = (x_{1i}, \dots, x_{mi})$ ,  $n \in \mathbb{N}$  is the number of sample points and  $m \in \mathbb{N}$  is the number of attributes, an interesting problem is to find  $z_k \in \mathbb{R}^m, k = 1, \dots, p$  that characterize the archetypal patterns in the data.

In archetypal analysis, the patterns  $z_1, \dots, z_p$  considered are mixtures of the data values  $x_i, i = 1, \dots, n$ . Furthermore, the only approximations to  $x_i, i = 1, \dots, n$  allowed are mixtures of the  $z_k, k = 1, \dots, p$ .

The value of  $p$  can be chosen.

1. If  $p = 1$ , one archetype is simply the average of sample points.
2. If  $1 < p < n$ , there is  $p$  archetypes on the boundary of the convex hull of  $x_i, i \in 1, \dots, n$ .
3. If  $p = n$ , every sample point creates an archetype.
4. If  $n < p$ , there is more archetypes than sample points in our segmentation.

For  $p = 1$ , the average does not need to equal to any sample point. This variant is inefficient for a market segmentation as we create only one general archetype that represents all sample points. On the other hand, this method is still useful – it is rather designed for mass marketing. We have to find at least two archetypes in order to find differences between sample points and segments of a given market.

For  $1 < p < n$ , this case is used for archetypal analysis. Usually, small integer values of  $p$  are chosen.

For  $p = n$ , there are as many archetypes as sample points. This is a very accurate method, because no bias occurs. This method is designed for one-to-one marketing. On the other hand, this method is useless for market segmentation: experience of previous researchers says that an efficient archetypal analysis has between 2 and 6 archetypes. The more archetypes, the longer time for computation is needed

and it is hard to propose marketing mix for every archetype of the archetypal analysis.

For  $n < p$ , the variant is clearly inefficient for market analysis.

Let  $X = (x_{ij}) \in \mathbb{R}^{n \times m}$  denotes a sample matrix,  $Z = (z_{kj}) \in \mathbb{R}^{p \times m}$  denotes an archetypal matrix,  $A = (\alpha_{ik}) \in \mathbb{R}^{n \times p}$  denotes coefficients of archetypes matrix and  $B = (\beta_{kj}) \in \mathbb{R}^{p \times n}$  denotes coefficients of data set. Let  $\alpha_i$ ,  $i = 1, \dots, n$  denotes  $i$ -th row of matrix A and  $\beta_k$ ,  $k = 1, \dots, p$  denotes  $k$ -th row of matrix B. The data matrix  $X$  is known from our survey, the matrices  $Z$ ,  $A$  and  $B$  are unknown and have to be computed. The value of  $m$  and  $n$  is determined given a survey and collected data. Matrix  $B$  is a sparse matrix, i.e. a matrix populated primarily with zeros; an archetype is a linear combination of just a few sample points.

For given  $x_1, \dots, x_n \in \mathbb{R}^m$ . The problem is to find  $z_1, \dots, z_p \in \mathbb{R}^m$  such that

$$z_k = \sum_{j=1}^n \beta_{kj} x_j, \quad (1)$$

and such that  $\alpha_{ik}$  and  $z_1, \dots, z_p$  minimize

$$SSR = \sum_{i=1}^n \left\| x_i - \sum_{k=1}^p \alpha_{ik} z_k \right\|^2 \quad (2)$$

subject to the following constraints:

1.  $0 \leq \alpha_{ik} \leq 1$  for  $i = 1, \dots, n$  and  $k = 1, \dots, p$ . Each  $\alpha_{ik}$  represents a value of each archetype in each sample. A value has to be non-negative, but the maximum level is 1. The larger  $\alpha_{ik}$ , the more the  $i$ -th observation is similar to the  $k$ -th archetype. If  $\alpha_{ik} = 1$ , the  $i$ -th observation will represent the  $k$ -th archetype. If  $\alpha_{ik} = 0$ , the  $i$ -th observation will have nothing in common with the  $k$ -th archetype.
2.  $\sum_{k=1}^p \alpha_{ik} = 1$  for  $i = 1, \dots, n$ . A sum of values of archetypes for each observation must equal to 1.

3.  $0 \leq \beta_{kj} \leq 1$  for  $k = 1, \dots, p$  and  $j = 1, \dots, n$ . Each  $\beta_{kj}$  represents a value of each sample point in each archetype. A value has to be non-negative, but the maximum level equals to 1. The larger  $\beta_{kj}$ , the more the  $k$ -th archetype is similar to the  $j$ -th sample point. If  $\beta_{kj} = 1$ , the  $k$ -th archetype will represent the  $j$ -th sample point. If  $\beta_{kj} = 0$ , the  $k$ -th archetype will have nothing in common with the  $j$ -th sample point.
4.  $\sum_{j=1}^n \beta_{kj} = 1$  for  $k = 1, \dots, p$ . A sum of values of sample points for each archetype must equal to 1.

The residual sum of squares in equation (2) can be rewritten as

$$SSR = \sum_{i=1}^n \left\| x_i - \sum_{k=1}^p \alpha_{ik} \sum_{j=1}^n \beta_{kj} x_j \right\|^2 \quad (3)$$

and the archetype problem is to find  $\alpha$ 's and  $\beta$ 's to minimize this SSR subject to the constraints 1. – 4. above.

Our goals are to find “the best”  $\alpha$  for given archetypes  $Z$  and to find “the best” archetypes  $Z$  for a given  $\alpha$ .

We use an iteration process, which consists of a repeating several optimization procedures in order to approach our target or result (see section 4). We alternate the optimization of entries of the matrix  $A$  and  $B$ .

However, we still do not know how many sample points consist of the  $k$ -th archetype,  $k = 1, \dots, p$  and which archetype prevails among the sample points. Hence, we use as new a modification of the Herfindahl index in archetypal analysis for a subsequent analysis. The Herfindahl index is defined as:

$$H = \sum_{i=1}^N s_i^2 \quad (4)$$

where  $s_i$ ,  $i = 1, \dots, N$ , is the market share of firm  $i$  in the market and  $N$  is the number of firms. The Herfindahl index is widely applied in competition law and antitrust. The higher the index, the more only

few companies control a given market sector. For  $H = 1$ , there is a monopoly, for  $H \rightarrow 0$ , there is a perfect competition in a given market sector. The square is used to give lower value to smaller numbers.

We would like to modify the Herfindahl index for our purposes as:

$$H = \sum_{k=1}^p \alpha_{ik}^2 \quad (5)$$

$i \in 1, \dots, n, H \in \langle \frac{1}{p}, 1 \rangle$ . It measures whether there is “a monopoly of archetype” or (rather) “a perfect competition of archetypes” in the  $i$ -th respondent. For  $H = \frac{1}{p}$ , all archetypes have the same value in a given respondent, for  $H = 1$ , the respondent is one of the archetypes.

We can also compute the average Herfindahl index for the given archetype  $k$ :

$$H_{A,A}(k) = \frac{1}{n} \sum_{i=1}^n \alpha_{ik}^2 \quad (6)$$

$k \in 1, \dots, p, H_{A,A}(k) \in \langle 0, 1 \rangle$ . The higher the  $H_{A,A}(k)$ , the larger average value has the  $k$ -th archetype in respondents. If a market strategy maker would like to choose only one archetype from  $p$  archetypes, we recommend the archetype  $k$  with the largest value of  $H_{A,A}(k)$ .

## 4 Overview of used algorithm in suitable software

In this section, we describe a procedure of a finding out archetypes from given data. The software we use was programmed by RNDr. Pavel Ranocha, Ph.D. Hence, we do not publish here the whole code. For the readers convenience, we describe only a pseudo-code of this software.

The algorithm consists of the following steps:

- 1) Data preparation:
  - 1.1) We have to create dummy variables as this method does not work with nominal values, but only with ordinal values.
  - 1.2) We have to standardize the data to unit variance, elsewhere a result will be dependent on the scale of each variable.
  - 1.3) We have to replace missing values, this method cannot work with missing data.
- 2) Recoding (see appendix A.1) and initialization.
- 3) Loop until SSR reduction is sufficiently small or the number of maximum iterations is reached:
  - 3.1) A starting set of archetypes  $Z$  is chosen as a random  $p$ -tuple of sample points. Hence, the matrix  $B$  is populated only with 0 or 1 and  $\sum_{j=1}^n \beta_{kj} = 1$  for  $k = 1, \dots, p$ . If  $\beta_{kj} = 1$ , the  $k$ -th archetype is the  $j$ -th sample point.
  - 3.2) Find the best  $\alpha$  for a given set of archetypes  $Z$ : solve  $n$  convex least square problems ( $i = 1, \dots, n$ )
$$\min_{\alpha_i} \frac{1}{2} \|x_i - \alpha_i Z\|^2 \quad \text{subject to } \alpha_i \geq 0 \quad \text{and} \quad \sum_{k=1}^p \alpha_{ik} = 1$$
  - 3.3) Recalculate archetypes  $\bar{Z}$ : solve system of linear equations  $X = \alpha \bar{Z}$

3.4) Find the best  $\beta$  for a given set of archetypes  $\bar{Z}$ : solve  $p$  convex least square problems ( $k = 1, \dots, p$ ).

$$\min_{\beta_k} \frac{1}{2} \|z_k - \beta_k X\|^2 \quad \text{subject to } \beta_k \geq 0 \quad \text{and} \quad \sum_{j=1}^n \beta_{kj} = 1$$

Go to 3.1 until the change is sufficiently small.

3.5) Recalculate archetypes  $Z$ :  $Z = X\beta$ .

3.6) Calculate sum of squares residuals SSR.

4) Post-processing.

For  $p > 1$ , the iteration process converges only to a local minimum and not to a desired global minimum. In order to be more sure about the correctness of our result, the used software repeats the iteration process until it get to the same result three times. Nevertheless, there is still no guarantee that the result is the global minimum.

## 5 Description of survey of students' behavior in banking sector

The aim of the survey was to analyze university students' attitude to money, saving, wealth and banking and financial products. The survey consisted of 27 questions, which were divided into 4 groups. 26 questions were obligatory, only one question was open and voluntary. There were demographical, factual questions and questions about attitude included in the survey.

**Group 1:** Questions included in this group were related to an extent of use of banking products.

1. *"How many bank accounts do you have?"* Respondents had to choose one answer out of following four possibilities: *0, 1, 2, 3 or more.*
2. *"Do you have any debit card?"* The definition: *"Debit cards are more common than credit cards. They are related to your bank account. When you pay with a debit card or withdraw money from an ATM with a debit card, you will use your own money. If you do not have money enough, a payment will not be possible."* followed the question as supposedly not everyone is familiar with the difference between a debit and a credit card<sup>3</sup>. Respondents could choose between *yes* or *no*.
3. *"Do you have any credit card?"* Also a definition: *"When you use a credit card, you will not spend your own money, but you borrow money from a banking institution. You do not have to have a bank account opened in the banking institution. A credit card is useful for payments in shops as you can use a period without interests. On the other hand, a withdrawal from an ATM is not*

---

<sup>3</sup>According to [14], 53% of respondents did not know the difference between a credit a debit card.

*useful*” was added. Respondents could choose between *yes* or *no*.

4. “*How many times a month do you pay with a card on average? When you do not have any card, please write 0.*”

5. “*How many times a month do you withdraw money from an ATM on average? When you do not have any debt card, please write 0.*”

6. “*How much do you withdraw from an ATM on average?*”

Question 4–6: Respondents had to fill in a non-negative number.

7. “*Do you use a standing order?*” Also definition: “*A regular sending of a well-defined amount of money to another bank account*” was added.

8. “*Do you use a collection?*” And definition: “*A permission given to a holder of another account to withdraw money from your account up to a limit, which you specify. Amounts of money can differ monthly, for example payments for a telephone*” followed.

9. “*Do you use Internet Banking?*”

Question 7–9: Respondents could choose between *yes* or *no*.

10. “*How many times did you visit a bank office during past year?*” Respondents had to fill in a non-negative number.

11. “*Do you have building savings contracted in your name?*” Respondents could choose from three possibilities: *Yes, I contracted in my own name./ Yes, my parents contracted it in my name./No.*

12. “Do you reach a maximum level of state support (do you save more than CZK 20 000 per year)?”
13. “Do you have a pension insurance?”

Question 12–13: Respondents could choose between *yes* or *no*.

**Group 2:** Questions included in this group were aimed to find out a level of respondents’ earnings and spending. We know that this type of questions is very delicate, therefore respondents had not have to answer in concrete amounts as intervals were proposed.

14. “How much do you spend on entertainment, food in restaurants, sports, telephone, culture,... (all except necessary expenses as food at home, living costs, etc.) per month?” Respondents could choose one interval out of five: CZK 0 – 1,000/ CZK 1,000 – 3,000/ CZK 3,000 – 5,000/ CZK 5,000 – 10,000/ 10,000 or more.
15. “How much do you spend on living, food, energy, etc. per month? (if your parents pay it, fill in 0).” Respondents had to fill in a non-negative number.
16. “How much is your monthly income?” Respondents could choose one interval out of five: CZK 0 – 3,000/ CZK 3,000 – 5,000/ CZK 5,000 – 10,000/ CZK 10,000 – 20,000/ CZK 20,000 or more.
17. “How much of your income do you not spend (do you save)?” Respondents could choose one interval from four ones: CZK 0/ CZK 1,000 or less/ CZK 3,000 or less/ CZK 3,000 or more.

**Group 3:** Questions in this group were not about hard facts, but about attitudes to money, savings and wealth.

18. *“To what extent do you agree with the following statement: I must permanently control, how much I spent in order not to fall in debts.”*
19. *“To what extent do you agree with the following statement: Every month I spent a little money, which I save.”*
20. *“To what extent do you agree with the following statement: I invest my spending in order to multiply them.”*
21. *“To what extent do you agree with the following statement: Money gives me an opportunity to do what I want.”*
22. *“To what extent do you agree with the following statement: I can easily give up most of things, which I have.”*
23. *“To what extent do you agree with the following statement: I buy everything, what I like.”*

Questions 18–23 wanted to know respondents’ attitude to money. Respondents could choose one interval out of five: *I strongly agree./ I agree./ I do not know./ I disagree./ I strongly disagree.*

**Group 4:** Questions in this group were demographical.

24. *“Your sex:”* Respondents had to choose between *male* or *female*. This question was included as we supposed that girls behave in other way than boys.
25. *“What school do you study?”* As there are many universities in the Czech Republic, schools were divided into seven groups: *IES FSV UK/ UK (Charles University in Prague), humanities/ UK,*

*medicine/ UK, natural sciences/ ČVUT (Czech Technical University) /VŠE (University of Economics in Prague) /other university.*

26. “*What grade do you attend?*” Respondents had to fill in a non-negative number that corresponds to their grade. Students of masters program should have included their years in a bachelors program. We supposed that respondent’s grade is strongly correlated with an age of respondent.
  
27. “*Do you use any other banking or financial products?*” An answering for this question was voluntary and open-ended.

The survey was created and maintained on a specialized web page `vyp1nto.cz`. Respondents could answer the questionnaire for three weeks, but most responses were obtained in the first two days. We have collected 471 responses, an average filling time was 5 minutes 15 seconds. Answers are available in appendix A.1. All data are available upon request.

## **Possible biases in our survey and remedies**

Biases in sample:

1. Wealthier students refused to answer.

We can not refuse this hypothesis. In order to have a high return rate (a ratio between filled-up and opened questionnaires), questions at which people could be more sensible (i.e. income, wealth) were not placed at the beginning of the questionnaire. Hence people, which have already spent some time by filling up the questionnaire, did not want to give up the filling up as they saw this type of questions. The survey began with “innocent” questions about an extent of usage of banking products. Demographic questions that could be boring were placed at the end of the questionnaire. And respondent did not have to fill in specific amounts; intervals

were usually offered. Finally, the return rate of our survey was at 81.2%. This means that almost one out of five individuals rejected answering or completing the survey as he/she saw the questions.

2. There were more younger students than older.

90% of the respondents studied between 1st and 3rd grade. For comparison, we also used a model with weighted respondents: each respondent got weight according to their grades – students of 1st grade: 0.658192, students of 2nd grade: 1.142156, students of 3rd grade: 0.850364 and student of master’s program: 2.33; and according to their university – students of IES FSV UK: 0.597436, students of UK excluding IES FSV UK: 0.751613, students of ČVUT: 1.350725, students of VŠE: 1.606897 and students of other universities: 1.579661. The aim of this weights is to provide every grade and every school (or a group of schools) with the same share in this model (25%, 20% respectively). Nevertheless, this weighed model still does not correspond to the real representation in population. If we used the real representation of schools, IES FSV UK would have a very little share among others. The weights have no impact on arising archetypes as the weights are used to data interpretation, not in counting archetypes.

Recall, the survey focused on attitudes rather than on specific amounts. The survey specialized on a quite small target group – students. According to [4], this group represents only 4% of the Czech population (the situation in 2012).

Biases in answers:

1. The question was badly understood.

To avoid this, we have included a wide range of questions. We have also included “soft” questions, which should support answers to “hard” questions, and explanation what do debit and credit card, a collection and a standing order mean. However, five respondents were excluded from the survey as their answers were

considered as inconsistent (e.g. one does not have any bank account, even though he/she has debit and credit card and uses Internet Banking or someone else does not have building savings, but he/she claimed he/she reached the full state support for building savings, etc.). Therefore, we included 466 respondents into our analysis.

## 2. Outliers can bias the results.

The results were recoded (see appendix A.1) – the answers were changed to numbers (e.g. “yes” was changed to 1, “no” to 0), because the numerical program work only with numbers. The questions that were evidently outliers (respondents most likely did not understand the question well) were changed to a median of answers to a given question and then recoded.

## 6 Analysis of the results and their interpretation

Using the software created by RNDr. Pavel Ranocha, Ph.D, we counted archetypes. The number of archetypes ranged from 1 to 6 archetypes. Unfortunately, the “elbow criterion” (see subsection 3.1) did not significantly help us to determine, what number of archetypes should be used as the SSE rate between 1 and 6 archetypes decreases with a constant pace, see A.2. We have decided to describe models with 1, 2 and 4 archetypes.

### 6.1 Model with 1 archetype – Average Respondent

If we use a variant with one archetype, the only archetype represents the mean of respondents included in our survey. According to the results of our survey, the average respondent has one bank account, he/she most likely (90%) has a debit card. On the other hand, with low probability (13%) he/she has credit card. With 28% probability the average respondent uses a standing order and with 20% probability he/she uses a collection. Almost everybody (88%) uses Internet Banking and visited a bank office on average 1.7 times last year. Most respondents (70%) have contracted building savings, 40% have a claim for full state support, only minority (15%) has contracted a pension insurance. Expenditures for entertainment, culture, sports, etc. of the average respondent per month amounts to CZK 1,000 – 3,000, necessary expenditures on living, energy, food at home amounts to less than CZK 5,000. Average monthly income is CZK 3,000 – 5,000. Savings do not reach CZK 1,000. Average respondent of our survey does not save much money, he/she does not invest. Money means for him/her the opportunity to do what he/she want. The representation of both sex is almost equal. The archetypal matrix for  $p = 1$  can be found in Appendix A.3.

Unfortunately, a student that fully corresponds to all of this criteria of the average respondent does not appear in our survey. Implying this, if a market strategy maker focuses on the average respondent, he/she will probably target only few people.

We can see that the extent of usage of banking products among college students exceeds the extent among general Czech population represented in [14]. One of the reasons of this phenomenon are the people included in the survey – young and with higher education. According to [14], people with higher education use more banking products. Also younger people use more products such as Internet Banking. In terms of savings, incomes and expenses, it is hard to compare their amount between college students and general Czech population as it is clear that a college student almost certainly has lower income than average Czech person.

## **6.2 Model with 2 archetypes**

In this model, two archetypes arose. According to our opinion, the archetypes differ mainly in the age of respondent and in an extent of usage banking products. Younger students use less of them in comparison with older students.

### **Freshman**

The “younger” means that typical representant of this archetype is between first and second grade. As has been stated above, the extent of usage of banking products is lower than older student’s one – even though the Freshman has almost one bank account and one debit card, he/she does not pay with it in a shop. He/she withdraws money from an ATM once a month and the amount of withdrawn money is less than CZK 500. He/she also has no credit card, does not use a standing order or a collection, but the Freshman usually uses Internet Banking. In terms of savings, the Freshman could have building savings, but hardly has a claim for full state support. He/she does not have any pension insurance. His/her expenditures and income are low as he/she most likely has no work and lives at home (or parents pay his/her necessary expenditures on behalf of her/him). Of course, the Freshman does not invest money and does not buy everything what he/she likes. This archetype represents both sexes.

## **Senior**

The “older” means that typical representant of this archetype is between second and third grade. The Senior has more bank accounts, one debit card and he/she can also have a credit card. The Senior pays with a card in shops. He/she withdraws from an ATM around CZK 500 once a week. The Senior uses a standing order and a collection rather than the Freshman. He/she also uses Internet Banking. The Senior has higher income and expenditures than the Freshman – spending between CZK 3,000 and 5,000 for entertainment, around CZK 5,000 for necessary expenditures and getting between CZK 5,000 and 10,000 a month. In terms of savings, the Senior does better than the Freshman – he has almost certainly contracted building savings, could have a claim for full state support. The Senior contracted a pension insurance. Money for the Senior means the possibility to do what he/she wants and he/she buys some things what he/she likes. This archetypes represents both sexes, but males prevail.

The Freshmen and the Seniors are almost equally represented in the sample. Unfortunately, the model with 2 archetypes does not tell us much about behavior as the archetypes arose based on age. Therefore, we would like to have a model with more archetypes, where behavior are more obvious.

### **6.3 Model with 4 archetypes**

The model with 4 archetypes seems to us as the best one as it represents the archetypes which arose based on behavior. The archetypes were accurately named as the Pauper, the Buffett, the Elvis and the Sleeping Beauty.

#### **Archetype 1 – Pauper**

Respondents representing Archetype 1 differ from the average respondent in a smaller extent of usage of banking products. The Pauper has no credit card, but he/she has a debit card, which is probably

left at home as the Pauper withdraws money from an ATM rarely, also seldom he/she pays with a card in a shop. A standing order or a collection do not belong into an enumeration of banking products that the Pauper uses. He/she most likely prefers payments in cash. However, it is not true that the Pauper does not know modern conveniences – the Pauper uses Internet Banking. The Pauper visits a bank office once every six months. In terms of savings, the Pauper does not differ from the average – has contracted building savings, maybe has a claim for full state support. The Pauper’s expenditures (dispensable or necessary) are not so high. It is most likely caused by the fact that his/her income is very low. The Pauper probably lives at home with his/her parents or parents pay the Pauper necessary expenditures. He/she has not contracted a pension insurance, yet, maybe because he/she are quite young and not worried about the future. In conclusion, the Pauper can not use all the banking products as he/she has low income and expenditures, but is aware of them. A total description of the archetype Pauper is available in Appendix A.5.

There is one respondent of the survey that fully corresponds to this archetype. Also the average value of the Pauper in a respondent (33%) is the highest among all four archetypes. And this not due to the fact that younger people have higher representation in the sample. If we use model with weighted respondents (see page 23), the average value of the Pauper in a respondent will be 30%. The number of respondents which consist of the Pauper from more than one half equals 139 (30% of respondents). On the other hand, almost one fifth of respondents (87 respondents) have nothing in common with the Pauper.

### **Archetype 2 – Buffett**

The Buffett has more than one bank account and also owns a debit card. On the other hand, the Buffett has not any credit card. In comparison with the Pauper, the Buffett pays with a card in a shop. The Buffett withdraws money from an ATM once a month and the amount of money is quite high. The Buffett is most likely aware that some withdrawals are not always for free or he/she is not afraid of

spending all this money at once. The Buffett also uses a standing order. On the other hand, the Buffett does not use a collection. Of course, the Buffett uses Internet Banking and visits bank office once in six months as the average respondent does. In comparison with the Pauper, dispensable expenditures are much higher, between CZK 3,000 and CZK 5,000. Expenditures for living, energy and food at home amount less than CZK 5,000. The Buffett can afford it as he/she has quite high monthly income, around CZK 10,000. Implying this, the Buffett can save more than CZK 3,000. This archetype is mainly represented by males. We can say that the Buffett is well aware about banking products and can use them as he has high expenditures and income. A total description of the archetype Buffett is available in Appendix A.5.

There are three respondents of the survey that fully correspond to this archetype. The average value of the Buffett in a respondent is lower than the Pauper's one – 24 percent. In the model with weighted respondents (see page 23), the average value is almost 30%. The number of respondents which consist of the Buffett from more than one half is only 77 (16.5% of respondents). 108 respondents (almost one fifth) have nothing in common with the Buffett.

### **Archetype 3 – Elvis**

The Elvis has more than one bank account and owns a debit card as the Buffett does. On the other hand, the Elvis differs from the others in credit cards. The Elvis can have a credit card (all other archetypes own no credit cards). This is reflected in expenditures. The Elvis has high expenditures for entertainment, sports, cultures, etc., even though the Elvis has not that high incomes (around CZK 5,000) as the Buffett has. Implying this, the Elvis spends all the money and has no savings. The Elvis also claims that he/she has to control his/her expenditure in order not to fall in debts. One of possibilities how to do it is to withdraw from an ATM smaller amounts several times per month. The Elvis states that he/she does it this way. This archetype uses Internet Banking and visits a bank office more times than the

Buffett does (the highest number of visits of all archetypes). However, the Elvis could have contracted building savings. Most likely his/her parents in behalf of him/her have done it in order to save him/her at least some money. The Elvis likes money and consumption, does not think about the future and prefer spending even saving money. A total description of the archetype Elvis is available in Appendix A.5.

There are two respondents that fully correspond to this archetype. The average value of the Elvis in a respondent is almost 30% (in the model with weighted respondents – see 2 on page 23, the average value is almost 29%). The number of respondents which consist of the Elvis from more than one half equals 104 (22 percent of respondents). 90 respondents have nothing in common with the Elvis.

#### **Archetype 4 – Sleeping Beauty**

Archetype 4 significantly differs from all three other archetypes. The Sleeping Beauty has neither bank account, debit nor credit card. Having said that, the Sleeping Beauty does not pay with a card in a shop, also does not withdraw money from an ATM, does not use Internet Banking. The Sleeping Beauty probably does not like banking products at all. The Sleeping Beauty also does not have building savings or a pension insurance. Their monthly expenditures for entertainment are quite low (around CZK 1,000), but higher than expenditures of the Pauper. The same is true for his/her monthly incomes. The Sleeping Beauty saves around CZK 1,000. This archetype most likely lives at home with his/her parents, or parents pay him/her necessary expenditures. As the Sleeping Beauty does not use any banking product, he/she also does not invest money. A total description of the archetype Sleeping Beauty is available in Appendix A.5.

There is one respondent that fully corresponds to this archetype. The average value of the Sleeping Beauty in a respondent is the lowest among all four archetypes – only 12.6% (in the model with weighted respondent – see 23, the value is 11%). Only 29 respondents consist of the Sleeping Beauty from more than one half. 36 percent of respondents have nothing in common with the Sleeping Beauty.

### Which archetype is the “best” one?

A market strategy maker will primarily focus on a specific archetype, which is close to an ideal consumer of his/her company, and hence it is attractive for a company. Thanks to archetypal analysis, the main differences between archetypes are clearly evident and can be easily compared.

If a given company would like to focus on the mass-represented archetype, the company will have to compare some indicators. First, we should compare  $H_{A,A}(k)$  in  $p$  archetype model. According to our survey, the highest  $H_{A,A}$  is for the Pauper (see table “Value of each archetype” on page 50). The Pauper will most likely gain higher incomes in the short term and subsequently use more banking products. On the second place, there is the Elvis as each respondent consists of this archetype from 30%, on average. The advantages of the Elvis incumbents in his/her consumer life. This archetype admits that consumption is preferred. In the future, this archetype will probably demand some loan for consumption, for houses (mortgages). Consequently, a banking institution can profit from this type of consumer. On the third place, there is the Buffett (respondents consist of this archetype from 24%, on average). In comparison with the Elvis, a banking institution will not probably profit from his/her loans. However, it can gain money from his/her investments’ fees. On the last place, there is the Sleeping Beauty as respondents consist of this archetype only from 12.6% on average. This archetype is not so suitable for a banking institution, because Sleeping Beauty does not use much banking products. Maybe it caused by his low age and income.

Archetypal analysis does not suggest us which archetype is the best one for our company (in the case that company is indifferent to types of consumer). However, it shows us the main differences between archetypes. The Herfindahl index modified for archetypal analysis shows an amount of representation of a given archetype. We can choose the most suitable archetype for our company and create a tailor-made advertisement targeting the specific archetype.

## 7 Summary

Market research is important for a company as it explores potential consumers. In order to meet consumers' needs, we can produce one general product, segment of the market, target each consumer or find a niche on a market.

Archetypal analysis was introduced in 1994 by Adele Cutler and Leo Breiman. This statistical method does not divide respondents into groups, but it finds out typical pure (or extremal) behavior. Every sample point is a mixture of some or all archetypes. Every archetype is a mixture of few sample points. The advantages of archetypal analysis is a very easy description and understanding of segments.

In the Czech Republic, there is no evidence of a research of students' behavior in banking sector, even though students will be important customers for a banking institutions as soon as they terminate their studies. Our survey explores the extent of usage of banking products, levels of earnings, expenditures and attitudes of college students at Czech universities.

In our archetypal analysis of students' behavior, we use models with 1, 2 and 4 archetypes. The only archetype in the model with 1 archetype represents the mean. The average respondent exceeds an average Czech person, in terms of using banking products. In the model with 2 archetypes, the archetype Freshman and the Senior arose based on the age of respondents.

In the model with 4 archetypes, the differences in behavior were more obvious. The Pauper has very low income, hence he does not use banking products so much, but is aware of them. The Buffett is well aware of banking products, has higher income, but does not spend all money. The Elvis prefers clearly consumption and has no savings. The Sleeping Beauty does not use any banking products, but has some income and expenditures.

We have introduced a modification of the Herfindahl index tailored for archetypal analysis. According to this measure, the archetype Pauper has the highest representation among our respondents. Also

the Elvis has quite high representation and most likely will be welcome by banking institutions as the Elvis will demand loan, mortgages, etc.

## References

- [1] P. Auger, T.M. Devinney, J.J. Louviere and P. Burke (2003): Consumer Assessment of Social Product Features: An Empirical Investigation Using Choice Experiments, *Working Paper*. <http://www.mbs.edu/index.cfm?objectid=951E3868-123F-A0D8-4261A8EB9D9E3479>
- [2] L.M. Borden, S. Lee, J. Serido, D.Collins (2007): Changing College Students' Financial Knowledge, Attitudes, and Behaviour through Seminar Participation *Journal of Family and Economic Issues*, **29(1)**. <http://www.usc.edu/dept/chepa/IDApays/publications/changingcollegestudents.pdf>
- [3] A. Cutler and L. Breiman (1994): Archetypal Analysis, *Technometrics* **36(4)**, 338-347.
- [4] Czech Statistical Office (2012): Statistical Yearbook of the Czech Republic 2011. <http://www.czso.cz/csu/2011edicniplan.nsf/engkapitola/0001-11-2010-2300>
- [5] M. Dvořák. *Metody shlukové analýzy a jejich aplikace v marketingu*, 2008. Master thesis, Charles University, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics. Supervisor: RNDr. Pavel Vaněček
- [6] M.J.A. Eugster and R. Leisch (2009): From Spider-Man to Hero - Archetypal Analysis in R, *Journal of Statistical Software* **30(8)**. <http://www.jstatsoft.org/v30/i08/paper>
- [7] B. Flury (2004): *A First Course in Multivariate Statistics*
- [8] I. Horáková, (1992): *Marketing v současné světové praxi*, Grada
- [9] G. Heinz, L.J. Peterson, R.W.Johnson and C.J.Kerk (2003): Exploring Relationships in Body Dimensions, *Journal of Statistics Education* **11(2)**.

<http://www.amstat.org/publications/JSE/v11n2/datasets.johnson.html>

- [10] T. Holub (2002): Credit usage and debt among college and university students, *ERIC digest*.  
<http://www.ericdigests.org/2003-2/credit.html>
- [11] Horáková (1992): *Marketing v současné světové praxi*, Grada
- [12] ING (2011): ING International Consumer Resourcefulness study.  
<https://www.ingfondy.cz/download/14097/2011%20Summary%20ING%20consumer%20resourcefulness%20study.pdf>
- [13] P. Kotler and G. Armstrong (2004): *Marketing*, Grada Publishing
- [14] Ministry of Finance of the Czech Republic (2011): Finanční gramotnost - Kvantitativní výzkum pro MF.  
<http://www.mfcr.cz>
- [15] Ministry of Finance of the Czech Republic (2011): Základní ukazatele vývoje penzijního připojištění v České republice ke dni 31. 12. 2011. <http://www.mfcr.cz>
- [16] Ministry of Finance of the Czech Republic (2011): Základní ukazatele vývoje stavebního spoření v České republice ke dni 31. 12. 2011. <http://www.mfcr.cz>
- [17] D. Peppers, M. Rogers and B.Dorf (1999): Is Your Company Ready for One-to-One Marketing? *Harvard Business Review*, January-February 1999.  
<http://hbr.org/1999/01/is-your-company-ready-for-one-to-one-marketing/ar/1>
- [18] D. Sharp, (2005): *Slovník základních pojmů psychologie*, Nakladatelství Tomáše Janečka
- [19] V. Šolc (2009): *Archetyp otce*, Triton

- [20] S. Li, P. Wang, J. Louviere and R. Carson (2003): Archetypal Analysis: A New Way to Segment Markets Based On Extreme Individuals. In *A Celebration of Ehrenberg and Bass: Marketing Knowledge, Discoveries and Contribution. Proceedings of the ANZMAC 2003 Conference, December 1-3, 2003* pages 1674-1679. <http://econ.ucsd.edu/rcarson/papers/ArchetypalANZMAC.pdf>

## A Appendix

### A.1 Results of the survey and recoding

| Question  | Answers         | Share | Rec. | Share |
|---|-----------------|-------|------|-------|
| 1. How many bank accounts do you have?              | 0               | 5.2   | 0    | 5.2   |
|   | 1               | 73.1  | 1    | 73.1  |
|   | 2               | 15    | 2    | 15    |
|   | 3+              | 6.6   | 3    | 6.6   |
| 2. Do you have any debit card?                      | yes             | 90.1  | 1    | 90.1  |
|   | no              | 9.8   | 0    | 9.8   |
| 3. Do you have any credit card?                     | yes             | 13.1  | 1    | 13.1  |
|   | no              | 86.9  | 0    | 86.9  |
| 4. How many times a month do you pay with a card?   | 0-2             | 32.6  | 0    | 32.6  |
|   | 3-5, 300+*      | 31.3  | 1    | 31.3  |
|   | 6+              | 36.1  | 2    | 36.1  |
| 5. How many times a month do you withdraw from ATM? | 0-1             | 29.6  | 0    | 29.6  |
|   | 2-4, 20+*       | 47.2  | 1    | 47.2  |
|   | 5+              | 23.2  | 2    | 23.2  |
| 6. How much do you withdraw from an ATM on average? | 0               | 11.8  | 0    | 11.8  |
|   | CZK 100 - 500   | 38.8  | 1    | 38.8  |
|   | CZK 501 - 1,000 | 32.6  | 2    | 32.6  |
|   | CZK 1000+       | 16.7  | 3    | 16.7  |
| 7. Do you use a standing order?                     | yes             | 30    | 1    | 30    |
|   | no              | 70    | 0    | 70    |
| 8. Do you use a collection?                         | yes             | 20.8  | 1    | 20.8  |
|   | no              | 79.2  | 0    | 79.2  |
| 9. Do you use Internet Banking?                     | yes             | 88.8  | 1    | 88.8  |
|   | no              | 11.2  | 0    | 11.2  |

\*Outliers were recoded to a median of answers of a given question:  
 1) Answers higher than 300 were changed to the median of 5, in question 4)  
 2) Answers higher than 20 were changed to the median of 3, in question 5.

| Question  | Answers                        | Share | Rec. | Share |
|---|--------------------------------|-------|------|-------|
| 10. How many times did you visit a bank office during past year?                        | 0                              | 17.6  | 0    | 17.6  |
|   | 1                              | 24.5  | 1    | 24.5  |
|   | 2                              | 27.9  | 2    | 27.9  |
|   | 3+                             | 30    | 3    | 30    |
| 11. Do you have building savings?   | Yes, I contracted it           | 5.2   | 1    | 70.6  |
|   | Yes, my parents contracted it. | 65.5  |      |       |
|   | No                             | 29.4  | 0    | 29.4  |
| 12. Do you reach max. level of state support?   | Yes                            | 44.2  | 1    | 44.2  |
|   | No                             | 55.8  | 0    | 55.8  |
| 13. Do you have a pension insurance?  | Yes                            | 15.7  | 1    | 15.7  |
|   | No                             | 84.3  | 0    | 84.3  |
| 14. How much do you spend on entertainment, food in restaurants, sports,etc. per month? | CZK 0-1,000                    | 18.7  | 0    | 18.7  |
|   | CZK 1,000 - 3,000              | 47.9  | 1    | 47.9  |
|   | CZK 3,000 - 5,000              | 24.7  | 2    | 24.7  |
|   | CZK 5,000                      | 7.5   | 3    | 8.8   |
|   | CZK 10,000+                    | 1.3   |      |       |
| 15. How much do you spend on living, food, energy, ... per month?                       | 0                              | 62.4  | 0    | 62.4  |
|   | CZK 1 - 5,000                  | 27.3  | 1    | 27.3  |
|   | CZK 5,001+                     | 10.3  | 2    | 10.3  |
| 16. How much is your monthly income?  | CZK 0 - 3,000                  | 34.1  | 0    | 34.1  |
|   | CZK 3,000 - 5,000              | 26    | 1    | 26    |
|   | CZK 5,000 - 10,000             | 29.2  | 2    | 29.2  |
|   | CZK 10,000 - 20,000            | 8.4   | 3    | 10.7  |
|   | CZK 20,000 or more             | 2.4   |      |       |

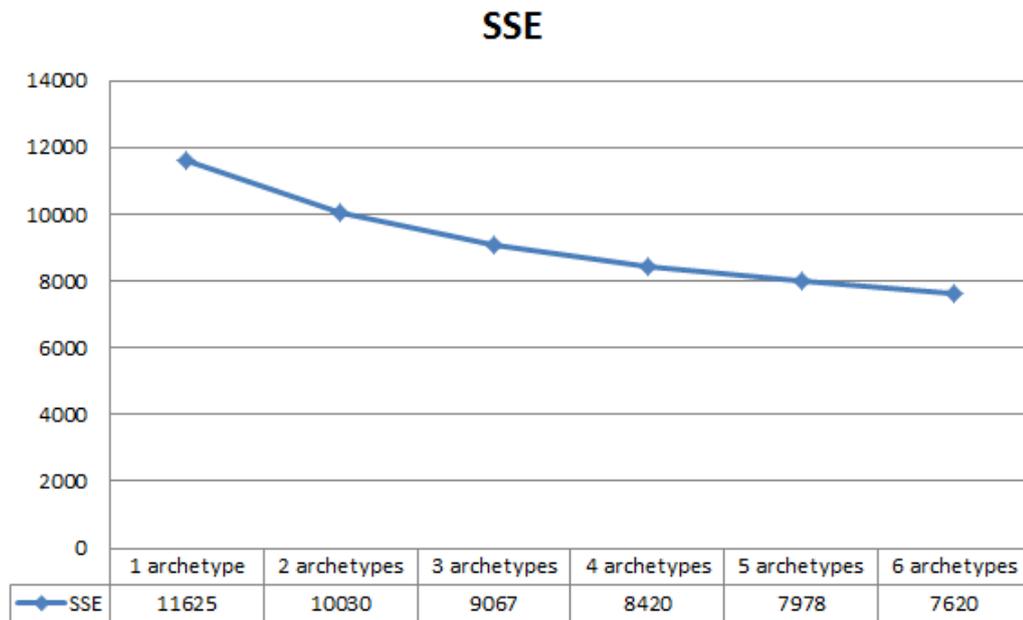
| Question  | Answers              | Share | Rec. | Share |
|---|----------------------|-------|------|-------|
| 17. How much of your income do you not spend?               | CZK 0                | 21.6  | 0    | 21.6  |
|   | CZK 1,000 or less    | 47    | 1    | 47    |
|   | CZK 1,000 - 3,000    | 20.6  | 2    | 20.6  |
|   | CZK 3,000 or more    | 10.9  | 3    | 10.9  |
| 18. I must permanently control, how much I spend.           | I strongly disagree. | 28.3  | 0    | 28.3  |
|   | I disagree.          | 31.1  | 1    | 31.1  |
|   | I do not know.       | 4.9   | 2    | 4.9   |
|   | I agree.             | 24.7  | 3    | 24.7  |
|   | I strongly agree.    | 10.9  | 4    | 10.9  |
| 19. Every month I do not spent a little money which I save. | I strongly disagree. | 23.8  | 0    | 23.8  |
|   | I disagree.          | 25.1  | 1    | 25.1  |
|   | I do not know.       | 6.4   | 2    | 6.4   |
|   | I agree.             | 23.4  | 3    | 23.4  |
|   | I strongly agree.    | 21.2  | 4    | 21.2  |
| 20. I invest my spending in order to multiply them.         | I strongly disagree. | 53.4  | 0    | 53.4  |
|   | I disagree.          | 22.5  | 1    | 22.5  |
|   | I do not know.       | 4.3   | 2    | 4.3   |
|   | I agree.             | 11.2  | 3    | 11.2  |
|   | I strongly agree.    | 8.6   | 4    | 8.6   |
| 21. Money gives me opportunity to do, what I want.          | I strongly disagree. | 5.2   | 0    | 5.2   |
|   | I disagree.          | 9.9   | 1    | 9.9   |
|   | I do not know.       | 7.7   | 2    | 7.7   |
|   | I agree.             | 40.1  | 3    | 40.1  |
|   | I strongly agree.    | 37.1  | 4    | 37.1  |

| Question  | Answers              | Share | Rec. | Share |
|---|----------------------|-------|------|-------|
| 22. I can easily give up most of things which I have. | I strongly disagree. | 15.7  | 0    | 15.7  |
|   | I disagree.          | 45.3  | 1    | 45.3  |
|   | I do not know.       | 16.7  | 2    | 16.7  |
|   | I agree.             | 18    | 3    | 18    |
|   | I strongly agree.    | 4.3   | 4    | 4.3   |
| 23. I buy everything, what I like.                    | I strongly disagree. | 16.1  | 0    | 16.1  |
|   | I disagree.          | 42.5  | 1    | 42.5  |
|   | I do not know.       | 7.9   | 2    | 7.9   |
|   | I agree.             | 29.2  | 3    | 29.2  |
|   | I strongly agree.    | 4.3   | 4    | 4.3   |
| 24. Your sex:   | Male                 | 51.7  | 0    | 51.7  |
|   | Female               | 48.3  | 1    | 48.3  |
| 25. What school do you study?                         | IES FSV UK           | 33.5  | *    | 33.5  |
|   | UK, humanities       | 21.5  | *    | 26.6  |
|   | UK, natural sciences | 3.2   |      |       |
|   | UK, medicine         | 2.0   |      |       |
|   | ČVUT                 | 14.8  | *    | 14.8  |
|   | VŠE                  | 12.4  | *    | 12.4  |
|   | other                | 12.6  | *    | 12.6  |
| 26. What grade do you attend?                         | 1                    | 38    | 1    | 38    |
|   | 2                    | 21.9  | 2    | 21.9  |
|   | 3                    | 29.4  | 3    | 29.4  |
|   | 4                    | 7.9   | 4    | 10.7  |
|   | 5                    | 2.6   |      |       |
|   | 6                    | 0.2   |      |       |

\*Attribute school served as a weight of respondent in the model with weighted respondent. See 2 on page 23

## A.2 SSE

Figure 1: SSE according to number of archetypes



### A.3 Model with 1 archetype

Archetypal matrix **Z**

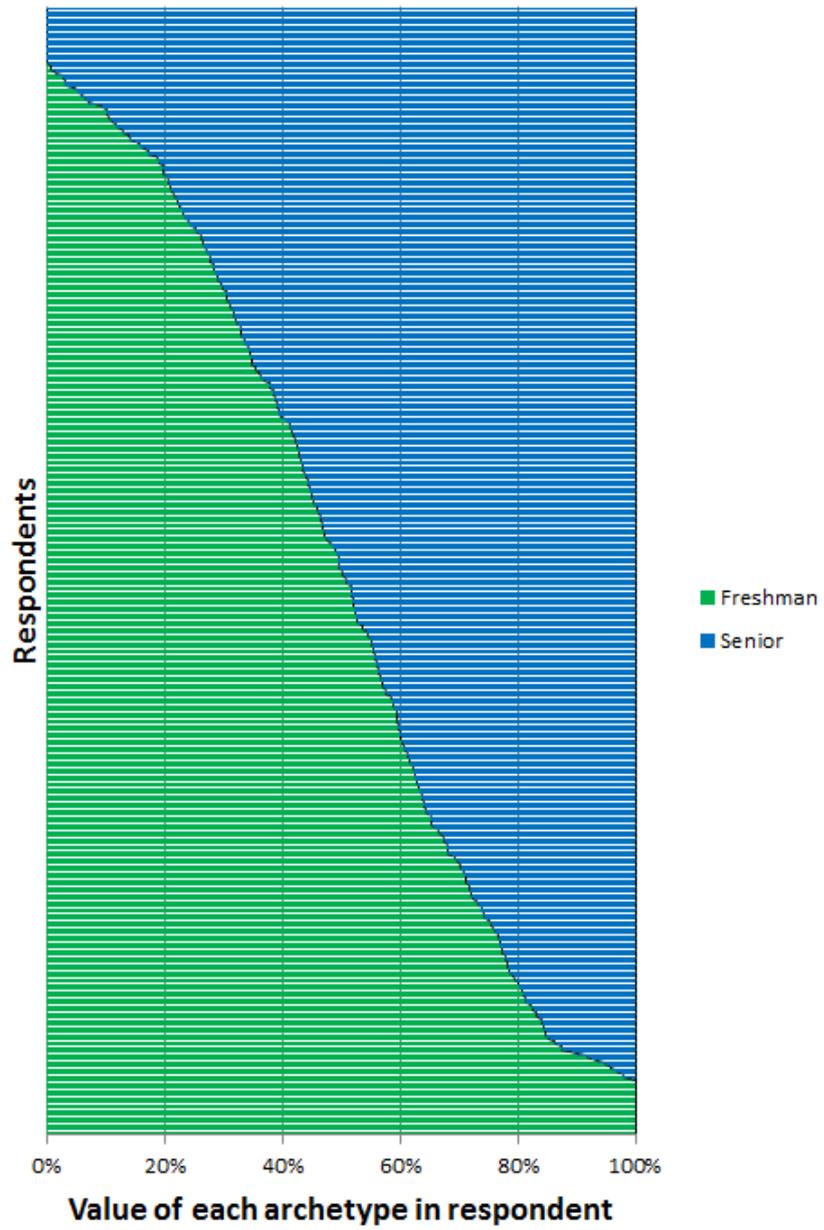
| Variable          | Archetype 1 |
|-------------------|-------------|
| Bank accounts     | 1.2318      |
| Debit card        | 0.9013      |
| Credit card       | 0.1309      |
| Pay with card     | 1.0343      |
| ATM               | 0.9356      |
| Amount from ATM   | 1.5429      |
| Standing order    | 0.2897      |
| Collection        | 0.2082      |
| Internet Banking  | 0.8884      |
| Bank office       | 1.7039      |
| Building savings  | 0.7060      |
| Max.state support | 0.4421      |
| Pension insurance | 0.1567      |
| Expens.staple     | 1.2361      |
| Expens.necces     | 0.4785      |
| Income            | 1.1652      |
| Not spent         | 1.2103      |
| Control money     | 1.5880      |
| Spend             | 1.9313      |
| Invest            | 0.9893      |
| Do what I want    | 2.9421      |
| Give up things    | 1.5000      |
| Buy everything    | 1.6309      |
| Sex               | 0.4828      |
| Grade             | 2.1288      |

## A.4 Model with 2 archetypes

### Archetypal matrix Z

|                   | Freshman | Senior |
|-------------------|----------|--------|
| Bank accounts     | 0.7384   | 1.7585 |
| Debit card        | 0.7295   | 1.0000 |
| Credit card       | 0.0199   | 0.2312 |
| Pay with card     | 0.1672   | 1.8850 |
| ATM               | 0.3692   | 1.4942 |
| Amount from ATM   | 0.7030   | 2.3716 |
| Standing order    | 0.0078   | 0.5807 |
| Collection        | 0.0344   | 0.3792 |
| Internet Banking  | 0.7208   | 1.0000 |
| Bank office       | 0.9793   | 2.4139 |
| Building savings  | 0.5282   | 0.8784 |
| Max.state support | 0.2356   | 0.6460 |
| Pension insurance | 0.0000   | 0.3111 |
| Expens.staple     | 0.2502   | 2.2013 |
| Expens.necces     | 0.0000   | 0.9945 |
| Income            | 0.0000   | 2.3992 |
| Not spent         | 0.6003   | 1.8265 |
| Control money     | 1.9787   | 1.2001 |
| Spend             | 1.5485   | 2.2823 |
| Invest            | 0.2279   | 1.7613 |
| Do what I want    | 2.2838   | 3.5865 |
| Give up things    | 1.6504   | 1.3407 |
| Buy everything    | 0.9941   | 2.2618 |
| Sex               | 0.5462   | 0.4159 |
| Grade             | 1.4035   | 2.8489 |

Figure 2: Value of each archetype in respondent



## A.5 Model with 4 archetypes

### Archetypal matrix $Z$

| Question | Archetype 1 | Archetype 2 | Archetype 3 | Archetype 4 |
|----------|-------------|-------------|-------------|-------------|
| 1        | 1.0261      | 1.7073      | 1.6228      | 0.0158      |
| 2        | 1.0000      | 1.0000      | 1.0000      | 0.0000      |
| 3        | 0.0000      | 0.0593      | 0.4021      | 0.0158      |
| 4        | 0.3011      | 1.4078      | 1.9120      | 0.0103      |
| 5        | 0.5268      | 0.9032      | 1.7513      | 0.0000      |
| 6        | 1.3112      | 2.4601      | 1.7163      | 0.0055      |
| 7        | 0.0322      | 0.8756      | 0.2471      | 0.0000      |
| 8        | 0.1866      | 0.4147      | 0.1823      | 0.0000      |
| 9        | 1.0000      | 1.0000      | 0.9833      | 0.0158      |
| 10       | 1.5780      | 1.7198      | 2.5600      | 0.0158      |
| 11       | 0.8997      | 1.0000      | 0.3444      | 0.3644      |
| 12       | 0.4996      | 0.9415      | 0.1036      | 0.1129      |
| 13       | 0.0322      | 0.4966      | 0.0570      | 0.0000      |
| 14       | 0.0316      | 2.0585      | 2.1078      | 0.7423      |
| 15       | 0.0962      | 0.9888      | 0.7763      | 0.0000      |
| 16       | 0.0000      | 2.8533      | 1.6209      | 0.3086      |
| 17       | 0.8254      | 2.9466      | 0.1326      | 1.1666      |
| 18       | 1.2201      | 0.2309      | 3.1231      | 1.8614      |
| 19       | 2.2950      | 3.6276      | 0.1540      | 2.0007      |
| 20       | 0.7486      | 3.1125      | 0.0314      | 0.1066      |
| 21       | 1.9618      | 3.5292      | 3.5565      | 2.6154      |
| 22       | 1.7794      | 1.5895      | 0.9486      | 1.7571      |
| 23       | 0.7661      | 2.3373      | 2.1929      | 1.5262      |
| 24       | 0.5367      | 0.1388      | 0.6771      | 0.5572      |
| 26       | 1.3572      | 2.9568      | 2.4987      | 1.6999      |

### Archetype 1 – Pauper

| Variable                                 | Amount            |
|--|-------------------|
| Bank accounts                            | 1                 |
| Debit cards                              | 1                 |
| Credit cards                             | 0                 |
| Payments with a card in a shop (monthly) | Rarely            |
| Withdrawals from ATM (monthly)           | 0.5               |
| Amount of money from ATM (on average)    | Around CZK 500    |
| Standing order                           | No                |
| Collection                               | Rarely            |
| Internet Banking                         | Yes               |
| Times in a bank office last year         | 1-2               |
| Building savings                         | Yes               |
| State support for building savings       | Someone has       |
| Pension insurance                        | No                |
| Non-necessary expenditures (per month)   | CZK 0 - 1,000     |
| Necessary expenditures (per month)       | CZK 0             |
| Income (per month)                       | CZK 0 - 3,000     |
| Saving (per month)                       | CZK 1,000 or less |
| Control money to avoid debts             | Disagree          |
| Save money                               | Rather agree      |
| Invest money                             | Strongly disagree |
| Money - opportunity to do what I want    | Rather disagree   |
| Give up most of things                   | Rather disagree   |
| Buy everything what I like               | Disagree          |
| Sex                                      | Both              |
| Grade                                    | 1st               |

**Archetype 2 – Buffett**

| Variable                                 | Amount              |
|--|---------------------|
| Bank accounts                            | More than 1         |
| Debit cards                              | 1                   |
| Credit cards                             | 0                   |
| Payments with a card in a shop (monthly) | 3-5                 |
| Withdrawals from ATM (monthly)           | 2-3                 |
| Amount of money from ATM (on average)    | around CZK 1,000    |
| Standing order                           | Yes                 |
| Collection                               | Sometimes           |
| Internet Banking                         | Yes                 |
| Times in a bank office last year         | 1 - 2               |
| Building savings                         | Yes                 |
| State support for building savings       | Yes                 |
| Pension insurance                        | Someone has         |
| Non-necessary expenditures (per month)   | CZK 3,000 - 5,000   |
| Necessary expenditures (per month)       | less than CZK 5,000 |
| Income (per month)                       | Around CZK 10,000   |
| Saving (per month)                       | Around CZK 3,000    |
| Control money to avoid debts             | Strongly disagree   |
| Save money                               | Strongly agree      |
| Invest money                             | Agree               |
| Money - opportunity to do what I want    | Strongly agree      |
| Give up most of things                   | Rather disagree     |
| Buy everything what I like               | Agree               |
| Sex                                      | Male                |
| Grade                                    | 3rd                 |

**Archetype 3 – Elvis**

| Variable                                 | Amount            |
|--|-------------------|
| Bank accounts                            | More than 1       |
| Debit cards                              | 1                 |
| Credit cards                             | Someone has       |
| Payments with a card in a shop (monthly) | 6                 |
| Withdrawals from ATM (monthly)           | 4                 |
| Amount of money from ATM (on average)    | CZK 500           |
| Standing order                           | Rather not        |
| Collection                               | No                |
| Internet Banking                         | Yes               |
| Times in a bank office last year         | 2 - 3             |
| Building savings                         | Rather not        |
| State support for building savings       | Rarely            |
| Pension insurance                        | No                |
| Non-necessary expenditures (per month)   | CZK 3,000 - 5,000 |
| Necessary expenditures (per month)       | Around CZK 5,000  |
| Income (per month)                       | Around CZK 5,000  |
| Saving (per month)                       | CZK 0             |
| Control money to avoid debts             | Agree             |
| Save money                               | Strongly disagree |
| Invest money                             | Strongly disagree |
| Money - opportunity to do what I want    | Strongly agree    |
| Give up most of things                   | Strongly disagree |
| Buy everything what I like               | Rather agree      |
| Sex                                      | Rather female     |
| Grade                                    | 2nd – 3 rd        |

#### Archetype 4 – Sleeping Beauty

| Variable                                 | Amount            |
|--|-------------------|
| Bank accounts                            | 0                 |
| Debit cards                              | 0                 |
| Credit cards                             | 0                 |
| Payments with a card in a shop (monthly) | 0                 |
| Withdrawals from ATM (monthly)           | 0                 |
| Amount of money from ATM (on average)    | 0                 |
| Standing order                           | No                |
| Collection                               | No                |
| Internet Banking                         | No                |
| Times in a bank office last year         | 0                 |
| Building savings                         | Rarely            |
| State support for building savings       | Rarely            |
| Pension insurance                        | No                |
| Non-necessary expenditures (per month)   | CZK 0 - 1,000     |
| Necessary expenditures (per month)       | CZK 0             |
| Income (per month)                       | CZK 0 - 3,000     |
| Saving (per month)                       | CZK 1,000 or less |
| Control money to avoid debts             | Rather disagree   |
| Save money                               | Rather agree      |
| Invest money                             | Strongly disagree |
| Money - opportunity to do what I want    | Agree             |
| Give up most of things                   | Rather disagree   |
| Buy everything what I like               | Rather disagree   |
| Sex                                      | Both              |
| Grade                                    | 1st – 3rd         |

**Alpha matrix - Selected rows**

| ID | Pauper | Buffett | Elvis  | Beauty | Sum | HI   |
|----|--------|---------|--------|--------|-----|------|
| 1  | 0.6222 | 0.3310  | 0.0000 | 0.0468 | 1   | 0.50 |
| 2  | 0.9409 | 0.0591  | 0.0000 | 0.0000 | 1   | 0.89 |
| 3  | 0.0424 | 0.0000  | 0.8076 | 0.1499 | 1   | 0.68 |
| 4  | 0.0000 | 0.1013  | 0.0000 | 0.8987 | 1   | 0.82 |
| 5  | 0.4555 | 0.1668  | 0.2001 | 0.1776 | 1   | 0.31 |
| 6  | 0.2253 | 0.6979  | 0.0000 | 0.0768 | 1   | 0.54 |
| 7  | 0.0000 | 0.3340  | 0.6660 | 0.0000 | 1   | 0.56 |
| 8  | 0.2944 | 0.6545  | 0.0004 | 0.0507 | 1   | 0.52 |
| 9  | 0.0000 | 0.1068  | 0.8932 | 0.0000 | 1   | 0.81 |
| 10 | 0.4267 | 0.4650  | 0.0000 | 0.1083 | 1   | 0.41 |
| 11 | 0.0000 | 0.0947  | 0.8226 | 0.0827 | 1   | 0.69 |
| 12 | 0.2351 | 0.0000  | 0.3927 | 0.3723 | 1   | 0.35 |
| 13 | 0.0008 | 0.1642  | 0.8349 | 0.0000 | 1   | 0.72 |
| 14 | 0.0000 | 0.0000  | 1.0000 | 0.0000 | 1   | 1.00 |
| 15 | 0.1711 | 0.7458  | 0.0831 | 0.0000 | 1   | 0.59 |
| 16 | 0.1411 | 0.2529  | 0.4694 | 0.1366 | 1   | 0.32 |
| 17 | 0.6786 | 0.0810  | 0.2298 | 0.0105 | 1   | 0.52 |
| 18 | 0.1309 | 0.4034  | 0.4656 | 0.0000 | 1   | 0.40 |
| 19 | 0.6427 | 0.0484  | 0.0569 | 0.2520 | 1   | 0.48 |
| 20 | 0.3239 | 0.3667  | 0.3094 | 0.0000 | 1   | 0.34 |

**Value of each archetype**

| Archetype                      | Pauper | Buffett | Elvis  | Sleeping beauty |
|--------------------------------|--------|---------|--------|-----------------|
| Average value                  | 0.3311 | 0.2479  | 0.2944 | 0.1266          |
| Model with weighted respondent | 0.3029 | 0.2960  | 0.2870 | 0.1114          |
| $H_{A,A}$                      | 0.1841 | 0.1200  | 0.1506 | 0.0619          |