Academic Board
Jacek Prokop (Chair), Warsaw School of Economics, Poland
Tomasz Gołębiowski, Warsaw School of Economics, Poland
Sándor Kerekes, Corvinus University of Budapest, Hungary
Bożena Leven, The College of New Jersey, USA
Rajiv Mehta, New Jersey Institute of Technology, USA
Mariusz Próchniak, Warsaw School of Economics, Poland
Katharine Rockett, University of Essex, Great Britain
Ryszard Rapacki, Warsaw School of Economics, Poland
Manuchehr Shahrokhi, Craig School of Business – California State University, Global Finance Association, USA
Jian Shi, Sichuan University, China
Meir Statman, Santa Clara University, USA
Theo Vermaelen, INSEAD, France
Christine Volkman, Wuppertal University, Germany
Charles V. Wait, Nelson Mandela Metropolitan University, South Africa
Marzanna Witek-Hajduk, Warsaw School of Economics, Poland
Ralf Zurbrugg, University of Adelaide, Australia

Editorial Board
Adam Szyszka – Editor in Chief
Krzysztof Falkowski – Vice Editor
Elżbieta Wąsowicz-Zaborek – Managing Editor

Honorary Editor in Chief
Jolanta Mazur

Statistical Editor
Piotr Zaborek

Text Editing and Proofreading
Kamila Grzesiak

Text editing and proofreading of English-language articles in the paper International Journal of Management and Economics are financed under contract 503 / P-DUN / 2017 from the resources of Ministry of Science and Higher Education allocated for activities which disseminate the science.
## Contents

Editorial ................................................................. 5  
*Adam Szyszka*

Asymmetric Shocks in the Euro Area: Convergence or Divergence? .............. 7  
*Dariusz K. Rosati*

Size, Value, and Momentum in Polish Equity Returns: Local or International Factors? ................................................................. 26  
*Adam Zaremba, Przemysław Konieczka*

Anchoring and Overconfidence: The Influence of Culture and Cognitive Abilities ................................................................. 48  
*Monika Czerwonka*

Culture-Based Rejection of Taboo-Infringing Imports ............................... 67  
*William D. Brice, Edward Chu, Anastasiya Brice*

Exploring Links Between Engaging Customers in Value Co-Creation and Product Innovativeness ................................................................. 82  
*Piotr Zaborek, Jolanta Mazur*

Identification of the Flow of Innovations in Tourism Related to Aesthetic Medicine ................................................................. 107  
*Eugenia Panfiluk, Aleksander M. Panasiuk, Elżbieta Szymańska*

Rectification ................................................................. 120
As the summer 2017 comes to an end, it is time to welcome our readers to the third issue of the *International Journal of Management and Economics* this year (volume 53). In it, all three disciplines of economic sciences are equally represented: we offer two papers on international economics, two on finance, and two on management. All of these papers are based on original empirical research grounded in a strong theoretical background. The papers in this issue, which cover the Eurozone, Poland, India, U.S., also generally provide international scope and context.

In the first paper, entitled “Asymmetric Shocks in the Euro Area: Convergence or Divergence?”, Dariusz K. Rosati discusses whether the Eurozone is an optimum currency area and the extent to which insufficient convergence between member countries may have contributed to particular problems. The author examines structural divergence in the Euro Area on the basis of the frequency and distribution of observed asymmetric shocks over the period 1996–2015. He argues that observed differences in the number of asymmetric shocks and their increased frequency during the crisis period may, to some extent, reflect differences in fiscal policy reactions by individual member countries, and do not necessarily indicate a widening structural divergence across the Eurozone. The author concludes that, while EA countries still differ structurally, the actual divergence between them may be smaller than what the number of revealed shocks would otherwise suggest.

The second article is on pricing factors in capital markets. Adam Zaremba and Przemysław Konieczka apply a number of cross-sectional tests to determine whether known pricing factors in the developed world markets exist also in the emerging market that is Poland. They find strong evidence for book-to-market and momentum effects, but only weak evidence for size premium. With the exception of the momentum factor, factors local to Poland are not correlated with their European and global counterparts. Additionally, the international value, size, and momentum factors perform poorly in explaining cross-sectional variation in Polish stock returns. These findings suggest that pricing factors are of a local character, and that employing internationally-based factor models to Poland's capital market is not a valid approach.

The third paper, by Monika Czerwonka, focuses on anchoring and overconfidence – two significant biases often mentioned in the behavioral finance literature. The study tests
whether cultural background and cognitive abilities play a role in a person’s susceptibility to these heuristic biases. It is based on a unique sample of students from Poland and India, and employs the Cognitive Reflection Test with open-ended questions.

The next paper, “Culture-Based Rejection of Taboo-Infringing Imports”, by William D. Brice, Edward Chu, and Anastasiya Brice, is a continuation of earlier research by Brice et al. (2016) presented in volume 50 of the IJME. This work further investigates the impact of culture on a market’s rejection of an imported product, and sheds more light on new aspects of international marketing of culture-laden products.

Piotr Zaborek and Jolanta Mazur explore links between engaging customers in value co-creation and product innovativeness. They find a significant positive correlation between the two, based on a representative sample of manufacturing and service firms. As a practical implication, they suggest that firms should intensify their efforts to engage customers in day-to-day operations to enhance innovativeness. However, the study indicates that not all aspects of co-creation provide a similar scale of positive effects – it seems that only the more involved procedures on the part of the company produce noticeable benefits.

In the final paper of this issue, Eugenia Panfiluk, Aleksander M. Panasiuk, and Elżbieta Szymańska investigate innovations in aesthetic medicine tourism in Poland. They identify innovation flow channels and describe the roles of the various parties involved, and argue that the innovative process in this tourism segment occurs in a complex multi-sectoral network. Medical service providers play the biggest role in innovation development, whereas accommodation and tourism service providers are responsible for diffusing the resulting innovations.

We hope that all of the papers in the current issue of the IJME will be a source of inspiration for our readers at the beginning of the new academic year 2017/2018.
Asymmetric Shocks in the Euro Area: Convergence or Divergence?

Abstract

The degree of structural divergence in the Euro Area is examined on the basis of the frequency and distribution of observed asymmetric shocks over the period 1996–2015. An asymmetric shock is defined as an opposite sign difference between the deviation of an individual country's GDP growth rate from a trend and the deviation of the EA-wide GDP growth rate from a trend. Two measures of asymmetric shocks are introduced, one based on exponential trend values and another on moving-average trend values. Geographical distribution of observed (“revealed”) shocks shows that EA member countries differ in terms of structural convergence, with a higher number of asymmetric shocks in countries that joined the EA at a later date. The distribution of asymmetric shocks over time shows two peaks in the number of shocks around 2002 and 2011, but no clear tendency towards more divergence is detected. As actual data may not provide a full picture of asymmetric shocks (given that countries with sufficient fiscal space could have neutralized their negative impact on GDP growth rates) a hypothesis on the existence of “non-revealed” negative asymmetric shocks is examined. Testing for correlation between public debt levels and GDP growth rate deviations confirms the existence of “non-revealed” asymmetric shocks in low-debt countries. In general, the observed differences in the number of asymmetric shocks in EA member countries (and their increases over time) may actually reflect different fiscal policy reactions in individual countries as well as the impact of financial and debt crises, and are not necessarily an indication of widening structural divergence across the EA.

Keywords: Euro Area, convergence, asymmetric shocks

JEL: F45
Introduction

The debt crisis in the Euro Area (EA) in 2010–2012 and subsequent efforts to reignite growth in indebted EA member countries have revived the discussion on whether the EA is an optimum currency area and to what extent insufficient convergence between member countries could have been a reason for the crisis. When the project to establish a currency union in Europe was launched in 1990s, it was widely recognized that prospective member countries differed significantly in terms of economic development levels, economic structure, as well as the stance of their fiscal and monetary policy. According to the theory of optimum currency areas [Mundell, 1961, 1969, 1973; McKinnon, 1963; De Graauwe, 2000], member countries of a successful currency union should display a sufficient degree of real convergence, structural convergence and nominal convergence, respectively. The lack of necessary convergence left prospective member countries vulnerable to country-specific – or asymmetric – shocks [Bayoumi, Eichengreen, 1992; Sørensen, Yoshia, 1998; Arreaza et al., 1998; Buti, Sapir, 1998; De Haan et al., 2007; Jonung, Drea, 2009; Estrada et al., 2012; Buti, Turrini, 2015]. Moreover, some economists warned that deepened regional specialization encouraged by the removal of trade and investment barriers and the introduction of a single currency would lead to more agglomeration of economic activities across the EA and more structural divergence across EA member countries, making them even more vulnerable to country-specific shocks [Krugman, 1991, 1993; Krugman, Venables, 1993].

Opposite views were also expressed. Some authors argued that the common currency and ensuing further deepening of trade and financial integration, as well as national policy coordination within the currency union, would accelerate convergence on all three fronts (real, structural and nominal) and gradually make member country economies more similar and less vulnerable to asymmetric shocks [European Commission, 1990; Frankel, Rose, 1998; Bayoumi, Eichengreen, 1992; Rose, 2000].

Empirical evidence on convergence in the EA is mixed, however. For instance, Mongelli and Wyplosz [2008] found that there has been a significant nominal convergence in the EA since the introduction of the euro, especially in terms of inflation levels and fiscal balances. They also show that income differentials between EA member countries have gradually declined. There is, however, less evidence on real and structural convergence. For instance, Estrada, Gali, López-Salido [2012] show that EA member countries generally converged until the financial crisis hit in 2008, and then diverged, especially with respect to labor markets and competitiveness levels. Also, Buti and Turrini [2015] provide evidence for strong real and structural convergence between 1999 and 2007.

One popular indicator of structural convergence between member countries of a currency union is the presence and frequency of asymmetric shocks. Some studies, using GDP growth rate differentials between member countries as a yardstick, suggest
that the number of asymmetric shocks in the EA actually increased, rather than decreased [IMF, 2013; Pisani-Ferry, 2012]. But simple GDP growth rates differentials are a rather crude measure of dispersion of individual growth rates, and may not be a good indicator of asymmetric shocks. Consequently, they may not be a good gauge of the degree of structural convergence or divergence. Thus, alternative measures of asymmetric shocks are needed. Two such measures are proposed in the paper.

An important underlying assumption for monetary integration is that asymmetric shocks in a currency union – if and when they occur – can and should be addressed by national fiscal policy. In the absence of national monetary policies, member countries should be able to sufficiently increase government spending and/or reduce taxes to stimulate the economy in times of negative demand and/or supply shocks. For this, a sufficiently large “fiscal space” in a national budget should be readily available. The concept of “fiscal space”, which derives from an empirical study by Bohn [1998], has been recently further developed by, among others, Ostry et al. [2010] and Ghosh et al. [2013].

In the EA context, member countries with low debt and ample fiscal space would be able – in principle – to react to negative shocks, thus preventing their GDP growth rates from falling. In those cases, asymmetric shocks cannot possibly be detected directly from drops in GDP growth rates because these potentially negative effects should have been effectively neutralized by active fiscal policy. We would call these shocks “non-revealed” asymmetric shocks. But in other EA member countries with very limited or no fiscal space (because of their high debt), fiscal policy could not be used to react to asymmetric shocks. In these cases, asymmetric shocks would have gone unabated, producing declines in their GDP growth rates. We would call these shocks “revealed” shocks. If the very existence of “non-revealed” shocks is confirmed for low-debt countries, the true number of asymmetric shocks in these countries may have actually been higher than measured by the “revealed” shocks only. This implies, first, that there may be less divergence between the high-debt and the low-debt countries, and second, that some “revealed” asymmetric shocks in high-debt countries may actually be symmetric shocks successfully neutralized in low-debt countries, but not neutralized in high-debt countries.

The main purpose of this paper is to assess to what extent the EA member countries have been subject to asymmetric shocks in the run-up to, and after the establishment of, the currency union, and to see whether the frequency of those shocks has been changing over time and across the EA. This will help answer the question whether EA member countries have generally converged or diverged since integration of their monetary policies. The other purpose, in this last context, is to check whether a higher number of revealed asymmetric shocks in individual member countries may have been related to the lack of fiscal shock-absorbing capacity in those countries.

The remaining part of the paper is organized into four sections. In the next section, the authors introduce measures of asymmetric shocks based on deviations of individual country GDP growth rates (from trend) of opposite sign to those of EA-wide growth rates.
deviations (from trend). The third section presents empirical evidence of the existence of asymmetric shocks in 19 EA member countries in 1996–2015. In the fourth section the relationship between the asymmetric shocks and the “fiscal space” is examined on the basis of public debt data, in order to check the hypothesis on existence of “non-revealed” asymmetric shocks. The last section states our conclusions.

**Alternative Measures of Asymmetric Shocks**

There is no established methodology of identifying and measuring asymmetric shocks. One possible approach is to take deviations of the actual, observed GDP growth rates from trend for individual EA member countries, and decompose these deviations into a common, EA-wide component and an individual country-specific component. The first component of the growth rates deviation from trend can be assumed to represent symmetric, EA-wide shocks while the second component may be assumed to represent asymmetric, country-specific shocks. This methodology has been applied, *inter alia*, by the International Monetary Fund [IMF, 2013].

The starting point is to estimate GDP exponential trend equations for individual countries and for the euro area as a whole.

\[ y_{i,t} = y_{i,0} \left(1 + \bar{r}_i\right)^t \quad i = 1, 2, \ldots, m; \quad t = 1, 2, \ldots, n \]

(1)

where \( y_{i,t} \) is the index of GDP level for country \( i \) in year \( t \), and \( \bar{r}_i \) is the average yearly rate of GDP growth for country \( i \). The next step is to calculate deviations of actual growth rates from theoretical trend values. These deviations are then decomposed into symmetric – or common – shocks, and asymmetric – or country-specific – shocks. Let these components be denoted as \( u_{EA,t} \) and \( u_{i,t} \), respectively, and the average GDP growth rate for the whole euro area as \( \bar{r}_{EA} \). Then, by definition, we have:

\[ r_{i,t} - \bar{r}_i = u_{EA,t} + u_{i,t} \]

(2)

where:

\[ u_{i,t} = r_{i,t} - \bar{r}_i - u_{EA,t} \]

(3)

and

\[ u_{EA,t} = r_{EA,t} - \bar{r}_{EA} \]

(4)

Combining (2), (3) and (4) we obtain:
Asymmetric Shocks in the Euro Area: Convergence or Divergence?

\[ u_{it} = \left( r_{it} - \bar{r}_i \right) - \left( r_{EA,t} - \bar{r}_{EA} \right) \] (5)

where \( u_{it} \) represents the country-specific component of the deviation of the GDP growth rate from trend for country \( i \) in period \( t \). This is how the asymmetric shock is defined in the IMF study. Equation (5) shows that these shocks can take positive or negative values.

However, defined this way, the IMF measure simply gauges the extent of dispersion in individual growth rates, rather than any real asymmetry in GDP performance. It can easily be demonstrated that not all asymmetric shocks identified with equation (5) are in fact asymmetric. For instance, if both expressions on the right-hand side of equation (5) are of the same sign and differ only in value, the IMF measure would signal an asymmetric shock. However, since the values of \( r_{it} \) (country-specific shock) and \( r_{EA,t} \) (EA-wide shock) are of the same sign, there is of course no “true” asymmetry, but only a difference in the respective growth rates.

Given this weakness of the IMF measure, the authors propose two alternative methods. The first alternative method starts again with taking deviations of actual growth rates from trend for individual countries and years, and comparing them with similar deviations for the EA as a whole. Then asymmetric shocks are defined as individual country growth rate deviations with signs opposite to those of EA deviations. For instance, if for a given year the actual growth rate for a given country was lower than the trend value for that country, while the actual EA growth rate was higher than the trend value for EA, the deviations are of the opposite sign, which means that the country was hit by an asymmetric shock. If, by contrast, the deviations are of the same sign (even if they differ strongly in size), there is no asymmetric shock. So, the definition requires that for each \( t \) and \( i \), one of the following conditions is strictly fulfilled:

\[ r_{it} - \bar{r}_i < 0 \land r_{EA,t} - \bar{r}_{EA} > 0 \] (6a)

or

\[ r_{it} - \bar{r}_i > 0 \land r_{EA,t} - \bar{r}_{EA} < 0 \] (6b)

The individual deviations from trend can, of course, be negative or positive, so we can have a negative asymmetric shock (6a) or a positive asymmetric shock (6b). The shocks also differ in size, reflecting different factors including fiscal policy stance, and have different impacts on GDP of individual countries. So, it was decided to distinguish between “weak” and “strong” shocks: if, for a given year, the absolute difference between the deviation from trend for a given country and the deviation from trend for the whole EA takes a value between 0 and 2 percentage points, it is called a “weak” asymmetric shock, and if this difference exceeds 2 percentage points, it is called a “strong” asymmetric shock. This measure will be defined as “Measure 1” of asymmetric shocks.
One important advantage of “Measure 1” is that, by contrast to the IMF measure, it is defined as an individual deviation with the opposite sign to that of the whole EA. This means that an asymmetric shock takes place only when the deviation from trend for an individual country is of different sign than the deviation from trend for the whole EA. Thus, “Measure 1” shows the “true” asymmetry in GDP changes.

The other alternative method is similar to “Measure 1” in that it defines asymmetric shocks as individual deviations with signs opposite to those of EA deviations. But contrary to “Measure 1” (and to the IMF measure), the individual deviations are this time calculated as differences between the actual GDP level index in a given member state for a given year, and the hypothetical trend level is estimated as a moving average of actual GDP index values. The advantage of using a moving-average smoothing model as opposed to the exponential trend model is that it allows one to eliminate the restrictive assumption about the constancy of the growth rate over time, and to take into account past, as well as future, observations to predict hypothetical values. At the same time, the moving-average model acts as a filter, allowing for smoothing away one-off shocks from the time series.

In order to calculate the deviations from the moving-average (MA) model for GDP levels, first a GDP level index is set in 1995 at 100, and then one obtains the series of hypothetical values for years 1996–2015 as five-year moving averages from actual GDP index values, according to the general formula:

\[
\hat{y}_t = \frac{\sum_{s=t-q}^{s=t+q} w_s y_s}{\sum_{s=t-q}^{s=t+q} w_s}
\]  

where \(\hat{y}_t\) is the hypothetical GDP index value in year \(t\), \(y_s\) is the actual GDP index value for year \(s\), with \(s\) taking values from \(t - q\) to \(t + q\), and \(w_s\) is the weight attached to \(y\) for year \(s\). Assuming \(q = 2\), and weights \(w_s\) equal to 1/5, equation (7) simplifies to (7a):

\[
\hat{y}_t = \sum_{s=t-2}^{s=t+2} w_s y_s
\]  

Next, the authors calculate the individual and EA-wide deviations of actual GDP index levels from the respective moving-average levels. Asymmetric shocks are defined as deviations of actual individual GDP index levels from the moving-average trend values with the opposite sign to the respective deviations of EA GDP index levels from the MA trend. So, the definition requires that for each \(t\) and \(i\), one of the following conditions is strictly fulfilled:

\[
y_{i,t} - \hat{y}_{i,t} < 0 \land y_{EA,t} - \hat{y}_{EA,t} > 0
\]  

where \(\hat{y}_{i,t}\) and \(\hat{y}_{EA,t}\) are the moving-average trend values for individual country and the EA, respectively.
or

\[ y_{i,t} - \hat{y}_{i,t} > 0 \land y_{EA,t} - \hat{y}_{EA,t} < 0 \] (8b)

Like in the previous case, one again distinguishes between “weak” shocks – taking values between 0 and 2 percentage points – and “strong” shocks – exceeding 2 percentage points. This measure will be defined as “Measure 2” of asymmetric shocks.

**Empirical Results**

We applied both measures of asymmetric shocks to a panel of data covering nineteen EA member countries (EA19) for the period 1996 to 2015. Figures for GDP growth rates were taken from Eurostat, except for growth rates for Luxemburg and Malta for the years 1996–2000, and Slovakia for the years 1996–1997, which were taken from the World Bank. The test starts with “measure 1” of asymmetric shocks. First, GDP trend equations were estimated for individual countries and for the EA19 as a whole. The average growth rates obtained and their standard deviations are shown in Table 1. Next, deviations of individual growth rates from trend and deviations of EA growth rates from trend were calculated, according to “measure 1”, in order to decompose the differentials and estimate country-specific GDP changes representing asymmetric shocks (according to equations (6) and (6a)). Out of 380 total observations, 87 asymmetric shocks were identified, of which 38 were “strong” asymmetric shocks. Graph 1 shows the distribution of asymmetric shocks across EA member countries, while Graph 2 shows the distribution of asymmetric shocks in EA countries over time. Both graphs show all asymmetric shocks and all “strong” asymmetric shocks.

| Table 1. Average trend values for GDP growth rates (in %) and standard deviations for the EA countries, 1996–2015 |
|---|---|---|
| Country | Average GDP growth rate, % | Standard deviation |
| BEL    | 1.67 | 1.4759 |
| DEU    | 1.26 | 2.0448 |
| CYP    | 1.92 | 3.0371 |
| EST    | 3.88 | 5.9495 |
| IRE    | 4.58 | 4.3275 |
| GRE    | 0.74 | 4.3836 |
| ESP    | 1.98 | 2.5102 |
| FRA    | 1.48 | 1.4887 |
| ITA    | 0.45 | 1.9742 |
### Table 1

<table>
<thead>
<tr>
<th>Country</th>
<th>Average GDP growth rate, %</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>LVA</td>
<td>3.89</td>
<td>5.9028</td>
</tr>
<tr>
<td>LTU</td>
<td>4.02</td>
<td>5.3205</td>
</tr>
<tr>
<td>LUX</td>
<td>3.45</td>
<td>3.3535</td>
</tr>
<tr>
<td>MLT</td>
<td>3.08</td>
<td>2.0920</td>
</tr>
<tr>
<td>NED</td>
<td>1.79</td>
<td>2.1340</td>
</tr>
<tr>
<td>AUT</td>
<td>1.69</td>
<td>1.6719</td>
</tr>
<tr>
<td>PRT</td>
<td>1.09</td>
<td>2.3516</td>
</tr>
<tr>
<td>SLV</td>
<td>2.39</td>
<td>3.2318</td>
</tr>
<tr>
<td>SVK</td>
<td>3.71</td>
<td>3.3554</td>
</tr>
<tr>
<td>FIN</td>
<td>2.01</td>
<td>3.2291</td>
</tr>
<tr>
<td>EA19</td>
<td>1.34</td>
<td>1.7961</td>
</tr>
</tbody>
</table>

**Source:** Eurostat, World Bank, own calculations.

---

**GRAPH 1.** “Measure 1” asymmetric shocks (AS) in the EA, by country, 1996–2015

As can be seen, the distribution of asymmetric shocks across EA countries is highly uneven. Italy did not register any shocks while Belgium, Germany, France, the Netherlands, Austria and Finland registered between 1 and 4 asymmetric shocks each, but no “strong” shocks. This group of countries shows the highest convergence with the EA as a whole. The low numbers of asymmetric shocks suggest that these countries most probably constitute parts of the optimum currency area on the basis of “Measure 1” criterion – they can be called the “core” countries. On the other end of the spectrum are Estonia, Greece, Latvia, Lithuania, Luxemburg, Malta and Slovakia, where asymmetric shocks have been much more
frequent (between 6 and 8 in each country), including many “strong” shocks (between 3 and 6). This group clearly shows much more divergence from the EA as a whole – this group can be called “a periphery”. However, it should be noted, that in case of Slovakia and the Baltic states most observed shocks took place before these countries joined the EA. The remaining countries (Cyprus, Ireland, Spain, Portugal and Slovenia) are somewhere in the middle – they have registered between 2 and 7 asymmetric shocks, but no more than 2 “strong” shocks each. It can be assumed that this group is probably a part of the optimum currency area, although to a lesser extent than the “core” group – they can be called a “semi-core” group.

The distribution of asymmetric shocks over time (Graph 2) shows an increase in the frequency of shocks in years 2002–2003 (20 shocks registered in two years, of which 9 were “strong” shocks), and in years 2010–2011 (19 shocks registered in two years, of which 11 were “strong” shocks). In the first episode, out of nine “strong” shocks, seven were observed in countries that were not at that time in the EA (The Baltics), and the two remaining shocks were registered in Greece. The second episode is unprecedented as it coincides with the sovereign debt crisis in the euro area in the aftermath of the global financial crisis. Out of eleven “strong” shocks in 2010–2011, seven were registered in EA member countries hardest hit by the financial crisis (the “program” countries – Greece, Ireland, Spain and Portugal) and three were registered in the Baltic countries that were not yet members of the EA. It would certainly be difficult to interpret these data as an indication of increasing structural divergence within the EA over time. The moderate increase in the number of “strong” asymmetric shocks over time is entirely due to the
sharp economic decline in the “program countries”. This is not, however, an indication of an increased structural divergence in the “old” EA (i.e. excluding the Baltics).

Next, we apply “Measure 2” of asymmetric shocks. Again, asymmetric shocks are defined as individual deviations with signs opposite to those of EA-wide deviations. But contrary to “measure 1” (and to the IMF measure), individual deviations are this time calculated as differences between the actual GDP level index in a given member state for a given year, and the hypothetical trend level estimated as a moving average of actual GDP index values. The number of asymmetric shocks identified with “Measure 2” is very similar to the number of asymmetric shocks identified with “Measure 1”. Of 380 total observations, there were 90 deviations with opposite sign, out of which 32 are “strong” negative asymmetric shocks. Graph 3 shows the distribution of negative asymmetric shocks across EA member countries, and Graph 4 shows the distribution of negative asymmetric shocks over time.


In terms of the geographical distribution of asymmetric shocks, Graph 3 conveys very much the same picture as Graph 1. Again, Italy experienced no shocks at all while Belgium, Germany, Spain, France, The Netherlands, Austria, Slovenia and Finland registered between 2 and 6 asymmetric shocks each (but no “strong” shocks) in the analyzed period. This is an indication of a high degree of structural convergence. This “core” group may again be assumed to constitute an optimum currency area on the basis of “Measure 2” criterion. On the other end of the spectrum are Estonia, Greece, Latvia, Lithuania and Slovakia, with 7 to 9 asymmetric shocks each, including between 3 and 8 “strong” shocks in each country. Again, each of these countries can be assumed to belong to the “periphery” group (but it should be noted, that most asymmetric shocks in the Baltic states and Slovakia took place before these countries joined the EA). Finally, Cyprus, Ireland, Luxemburg, Malta and Portugal fall somewhere in the middle of the range, with 2 to 6 asymmetric shocks
Asymmetric Shocks in the Euro Area: Convergence or Divergence?

(but only one “strong” shock) were registered in each of them. This is a “semi-core” group, according to “Measure 2”. The main differences between “Measure 1” and “Measure 2” classifications concern Spain (moved from “semi-core” to “core”), Luxemburg (moved from “periphery” to “semi-core”) and Slovenia (moved from “semi-core” to “core”). Otherwise, the results obtained from the two measures is very similar. A summary classification of member countries by both measures is given in Table 2.


Changes in the number of negative shocks over time (Graph 4) also show a higher number of asymmetric shocks in years 1998–2002 and 2010–2014, including “strong” shocks. This particular pattern may reflect the impact of the dot-com crisis during the first episode and of the financial-cum-debt crisis during the second episode. As far as the first episode is concerned, all ten “strong” shocks observed took place in countries that were not yet members of the EA (The Baltics and Slovakia). During the second episode, of fifteen “strong” shocks observed in 2010–2014, more than half (eight) took place in the Baltic states (Estonia joined the EA in 2011, and Latvia in 2014, and Lithuania in 2015), and five in the countries that were most indebted and in financial distress (Cyprus, Greece, Ireland and Portugal). Again, this distribution of asymmetric shocks over time does not provide sufficient evidence of more divergence in the EA. In fact, the observed – rather moderate – increase in the number of asymmetric shocks (both “weak” and “strong”) over time can be more than explained by the shocks taking place in the “program countries”.

Two conclusions can be drawn. First, EA member countries can be broadly classified into three categories. As shown in Table 2, the “core” group includes Belgium, Germany, France, Italy and the Netherlands (the EU founding member states minus Luxemburg) plus
Austria and Finland. The “semi-core” group includes Cyprus, Portugal, Luxemburg, Spain and Slovenia (“semi-core”). These two groups of countries display a high, or moderately high, degree of structural convergence, and can be assumed to constitute parts of an optimum currency area. The third group – the “periphery” – consists of the EA member countries with a generally much higher number of shocks – Estonia, Ireland, Greece, Latvia, Lithuania, Malta and Slovakia – where the degree of structural convergence with the EA is low, or relatively low. These countries cannot be assumed to be parts of the optimum currency area. However, given that most of these countries joined the EA only recently, they will probably converge in the future.

**TABLE 2. EA member countries, categorized by the degree of convergence measured by the frequency and strength of asymmetric shocks (AS)**

<table>
<thead>
<tr>
<th>EA member countries, by categories</th>
<th>“Measure 1”</th>
<th>“Measure 2”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of all AS</td>
<td>No. of “strong” AS</td>
</tr>
<tr>
<td>&quot;Core&quot; countries:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Germany</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>France</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Italy</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Austria</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Finland</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>&quot;Semi-core” countries:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyprus</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Ireland</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Spain</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Luxemburg</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Portugal</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Slovenia</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>&quot;Periphery” countries:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estonia</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Greece</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Latvia</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Lithuania</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Malta</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Slovakia</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Total number of AS</td>
<td>87</td>
<td>38</td>
</tr>
</tbody>
</table>

Source: own study.

Second, the data broadly show two episodes of a higher number of asymmetric shocks (including “strong” shocks) – in years 1998–2004 and in years 2010–2014. Almost all of these shocks were observed in countries that either were not yet EA members (the Baltics,
Slovakia), or were most hit by financial and economic crisis (Greece, Ireland, Cyprus, Portugal, Spain – the “program countries”). A moderate increase of shocks over time that can be detected from the data is fully consistent with what happened in these two groups of countries. The results obtained demonstrate that there was no general tendency towards more structural divergence among EA member countries in the analyzed period [see also Buti, Turrini, 2015]. It can also be observed that EA member countries that joined the currency union at a later date, and with a low initial degree of structural convergence, have also been more vulnerable to economic and financial crises.

**Asymmetric Shocks and the Fiscal Policy Stance**

Generally, when a negative shock hits a country would normally attempt to react to it with fiscal policy expansion, spending more or taxing less. The available space for fiscal expansion depends of course primarily on the level of the country’s indebtedness – the higher the debt in proportion to GDP, the lower the “fiscal space” available. In principle, a country with a sufficiently large fiscal space should therefore be able to compensate for the impact of the shock. In that case, GDP growth rate should remain broadly unaffected.

If, however, the country’s fiscal space is limited or nonexistent, then the adversely affects GDP growth rate. If these assumptions are correct as a general proposition, then the asymmetric shocks that were analyzed in the previous section should be considered as “revealed” asymmetric shocks, i.e. the shocks that did occur because they were not neutralized by shock-absorbing measures under national fiscal policy. So, it is plausible, that some shocks may not eventually be observed in GDP statistics precisely because their impact on GDP growth rates was compensated – partially or fully – by parallel expansionary fiscal policy. If this was the case, the number of actually observed – or “revealed” – shocks (symmetric or asymmetric) in countries with sufficiently large fiscal space could in fact have been lower relative to other countries where fiscal space was limited or nonexistent.

The “revealed” asymmetric shocks – as identified by “Measure 1” and “Measure 2” – may not therefore offer a full picture of the degree of structural convergence or divergence in the EA. In particular, if high-debt member countries with little (or no) fiscal space have registered more “revealed” asymmetric shocks, this did not necessarily indicate their stronger structural divergence. And *vice versa*, if the low-debt countries have registered less asymmetric shocks, this did indicate their stronger convergence. However, “non-revealed” shocks could not be directly observed because their impact on output was compensated by fiscal policy, leaving GDP growth rates broadly unchanged. If the existence of “non-revealed” shocks could be confirmed for low-debt member countries, this could indicate that the actual level of convergence between low-debt and high-debt EA member countries is higher than suggested by actual GDP growth data.
This possibility was examined by analyzing the relationship between “revealed” asymmetric shocks and the level of public debt. It is assumed that countries with high debt levels have limited or no fiscal space, while countries with low debt levels have sufficient fiscal space to absorb negative asymmetric shocks. The relationship between the debt level and the fiscal space is not linear, however. Public debt becomes an effectively binding constraint on fiscal policy only at relatively high debt levels. To properly control for the level of public debt, all 380 annual observations have therefore been divided into two subsets – observations with high debt levels and observations with low debt levels – with the threshold separating the two subsets set at 90% of GDP. In selecting this specific threshold value the authors draw on Reinhard and Rogoff [2010] and Cechetti et al. [2011], who found that the 90% threshold is the critical level above which public debt starts to negatively influence growth in developed economies. One assumes therefore that for high-debt observations (with debt levels of 90% of GDP or more) there would generally be no fiscal response to shocks, while for low-debt observations (with debt levels below 90% of GDP) fiscal policy would effectively neutralize asymmetric shocks. If the hypothesis on the existence of “non-revealed” asymmetric shocks is correct, the revealed negative asymmetric shocks should generally more often be observed in years characterized by higher debt levels. Table 3 shows the regression results for GDP growth rates deviations and public debt levels.

As can be seen, controlling for the debt level demonstrates that the relationship between public debt levels and GDP growth rate deviations is different for high-debt level observations and for low-debt level observations. For high-debt level observations, there is a statistically significant negative correlation between debt levels and GDP growth rates deviations (stronger for “Measure 1” shocks and weaker for “Measure 2” shocks, as shown by equations (10) and (14), respectively). This means that, for the given panel data, high levels of public debt tend to be associated with larger negative deviations of individual GDP growth rates from EA-wide deviations (because the structural coefficients are significant and negative). This implies that in high-debt member countries with little or no fiscal space, the revealed negative asymmetric shocks tend to be more frequent. On the other hand, for low-debt level observations (below 90% of GDP), no statistically significant relationship is generally observed between the debt level and deviations of GDP growth rates (as shown by equations (9) and (13)). This lack of correlation means that in member countries with abundant fiscal space negative asymmetric shocks may have been neutralized with autonomous fiscal policy reaction. This implies the existence of “non-revealed” asymmetric shocks.

The relationship between debt-levels and asymmetric shocks (individual GDP growth rates deviations of opposite sign to EA-wide deviations from trend) is less clear. For “Measure 1” shocks, a strong statistically significant negative correlation is observed between debt levels and GDP growth rate deviations for high-debt countries, as shown by
equation (12), which is consistent with previous results. However, no significant correlation has been found for “Measure 2” shocks (equation (16))

<table>
<thead>
<tr>
<th>Eq. No</th>
<th>Coverage</th>
<th>No. of obs.</th>
<th>Constant</th>
<th>Debt, b)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(9)</td>
<td>All observations for debt &lt; 90% SEE</td>
<td>301</td>
<td>1.3618 (0.4523)</td>
<td>0.0229 (0.0891)</td>
<td>0.0215</td>
</tr>
<tr>
<td>(10)</td>
<td>All observations for debt &gt; 90% SEE</td>
<td>79</td>
<td>7.3386 (1.7393)</td>
<td>−0.0685*** (0.0152)</td>
<td>0.2079</td>
</tr>
<tr>
<td>(11)</td>
<td>Asymmetric shocks for debt &lt; 90% SEE</td>
<td>72</td>
<td>0.4433 (0.5277)</td>
<td>−0.0191* (0.0109)</td>
<td>0.0419</td>
</tr>
<tr>
<td>(12)</td>
<td>Asymmetric shocks for debt &gt; 90% SEE</td>
<td>15</td>
<td>9.7937 (3.4922)</td>
<td>−0.0905** (0.0287)</td>
<td>0.4334</td>
</tr>
<tr>
<td>(13)</td>
<td>All observations for debt &lt; 90% SEE</td>
<td>301</td>
<td>0.5997 (0.5849)</td>
<td>0.0143 (0.0115)</td>
<td>0.0051</td>
</tr>
<tr>
<td>(14)</td>
<td>All observations for debt &gt; 90% SEE</td>
<td>79</td>
<td>4.1136 (1.8576)</td>
<td>−0.0368** (0.0163)</td>
<td>0.0624</td>
</tr>
<tr>
<td>(15)</td>
<td>Asymmetric shocks for debt &lt; 90% SEE</td>
<td>73</td>
<td>0.2418 (0.4058)</td>
<td>0.0008 (0.0084)</td>
<td>0.0001</td>
</tr>
<tr>
<td>(16)</td>
<td>Asymmetric shocks for debt &gt; 90% SEE</td>
<td>17</td>
<td>1.6282 (2.0656)</td>
<td>−0.0098 (0.0183)</td>
<td>0.0187</td>
</tr>
</tbody>
</table>

*Asymmetric shocks defined as individual countries’ GDP growth rate deviations with signs opposite to those of the EA-wide deviations; standard errors of estimate (SEE) in brackets; public debt levels as % of GDP.
Source: own calculations.

On balance, it can be argued that the results obtained suggest the existence of “non-revealed” shocks, both symmetric and asymmetric. There are two possible implications. First, the EA member countries may actually be less structurally different than suggested by the “revealed” shocks. This is especially true for the differences between the high-debt and the low-debt countries. The differences in the number of observed asymmetric shocks in the EA member countries may to some extent reflect different fiscal policy reactions in individual countries and are not necessarily an indication of widening structural divergence across the EA. Second, it is possible that some “revealed” asymmetric shocks in high-debt countries may actually be symmetric shocks that have been successfully neutralized by low-debt countries. This again would suggest less structural divergence across the EA than suggested by observed GDP growth rate deviations.
Conclusions

The main purpose of the paper has been to check for the occurrence of asymmetric shocks in EA member countries in the period 1996–2015, in order to assess the extent of structural divergence between EA member countries and across time. Another purpose was to check whether a higher number of revealed asymmetric shocks in some individual member countries may have been related to the lack of sufficient shock-absorbing capacity in their public finances – or a lack of sufficient fiscal space. Before identifying asymmetric shocks, one first critically assesses the standard definition of an asymmetric shock as a country-specific component of the deviation of an individual country GDP growth rate from its trend over the analyzed period. We argue that this measure simply gauges the extent of dispersion in individual growth rates, rather than any true asymmetry in GDP performance.

To properly identify asymmetric shocks, two alternative measures are proposed. “Measure 1” defines asymmetric shocks as individual countries’ growth rate deviations from trend with signs opposite to those of the EA growth rates deviations from its EA trend. “Measure 2” differs from “Measure 1” in that the individual deviations are this time calculated as differences between the actual GDP level index in a given member state for a given year, and the hypothetical trend level estimated as a moving average of actual GDP index values. We have also divided the asymmetric shocks into “weak” shocks (growth differentials of less than 2 percentage points) and “strong” shocks (growth differentials exceeding 2 percentage points).

Both measures were applied to the panel of data covering 19 EA member countries over the twenty year period of 1996–2015. With respect to geographical distribution of asymmetric shocks, there are three distinct groups of countries in the EA – the “core”, the “semi-core” and the “periphery”. In the “core” group, which includes Belgium, Germany, France, Italy, the Netherlands, Austria and Finland, no “strong” shocks and very few “weak” shocks were generally observed. The “semi-core” group, which includes Cyprus, Ireland, Luxemburg, Portugal, Spain and Slovenia, registered more shocks but very few “strong” shocks. These two groups of countries display a high, or moderately high, degree of structural convergence, and can be assumed to broadly constitute parts of an optimum currency area. In the “periphery” group, which includes Estonia, Greece, Ireland, Latvia, Lithuania, Malta and Slovakia, the frequency of asymmetric shocks is much higher, which suggests that these countries may not belong to the same optimum currency area as the first group of countries. However, given that most of the “periphery” countries joined the EA only recently, they may still converge in the future. With respect to the distribution of asymmetric shocks over time, we found a moderate increase in the frequency of shocks in years 1999–2000, and a stronger increase in years 2008–2012. This increase, however, reflects a higher number of shocks observed in the “program” countries most affected by
the financial-cum-debt crisis, and therefore should not be interpreted as a symptom of increased structural divergence. It can be noted, that EA member countries with a low initial degree of structural convergence, as well as new EA members, have also been most vulnerable to economic and financial crises.

Finally, we addressed the question to what extent negative asymmetric shocks have materialized because they were not – or could not be – neutralized by an active national fiscal policy response due to the lack of necessary fiscal space. We tested the hypothesis on the possible existence of “non-revealed” shocks in low-debt countries, i.e. shocks that could not have been observed because they were effectively neutralized by countercyclical fiscal policy. We also examined the relationship between revealed asymmetric shocks and public debt levels in individual countries. All observations were divided into two subsets – low-debt observations (with public debt levels below 90% of GDP) and high-debt observations (with public debt at or above 90% of GDP). We found that there is a statistically significant negative correlation between high-debt observations and GDP growth rate deviations, and much lower or no correlation between low-debt observations and GDP growth rate deviations. This implies the existence of “non-revealed” asymmetric shocks in low-debt countries.

To sum up, while EA countries still differ structurally, the actual divergence may be smaller than suggested by the number of revealed shocks. Rather, the observed differences in the number of asymmetric shocks in EA member countries (and their increase in the crisis years) may actually reflect different fiscal policy reactions in individual countries as well as the impact of financial and debt crises, and are not necessarily an indication of widening structural divergence across the EA.

Notes

1 Author’s email address: drosat@sgh.waw.pl
2 The issue is of practical importance. The idea to establish a “fiscal capacity” for the Eurozone, put forward, inter alia, in the so-called “Five Presidents Report”, is based on the assumption that the scope for asymmetric shocks in the EA is still so large that a common budget is needed to provide a public risk-sharing mechanism that would help smooth the negative impact of shocks across member countries [European Commission, 2015].
3 The two-percentage points limit represents 0.63 of standard deviation of all GDP growth rate differentials.
4 The two-percentage point limit represents 0.55 of standard deviation of all GDP growth rate differentials.
5 The rationale for including the „pre-euro” years (1996–1998) is that in the run-up to the euro all prospective member countries had largely fixed exchange rates and had to comply with nominal
convergence criteria, exactly as if they were already members of the currency union. This is also the reason for including member countries that joined the Eurozone at later dates, as they similarly had fixed exchange rate regimes in their pre-euro years.

It should be remembered that the smaller number of asymmetric shocks for the biggest EA member countries (Germany, France, Italy) may to some extent reflect their larger share in the EA’s GDP, and therefore relatively smaller deviations of individual GDP growth rates from EA-wide values.

Luxemburg presents a special case, with quite a high number of shocks (especially according to “Measure 1”), even though this country is obviously very much integrated with the rest of the EA through trade and investment. One possible explanation for this anomaly is a very high share of financial services in GDP (the average share for 1996–2015 is 25.8%, compared with the average for the whole EA of 5.0%). The specific structure of Luxemburg’s economy makes it five time more vulnerable to shocks originating in the financial services sector than the rest of the EA, but much less so to shocks originating in manufacturing. While it can be argued that Luxemburg is a unique case, for the purpose of the paper it has been classified as “semi-core” rather than “periphery”.

The concept of a “fiscal space” can be defined as the distance between the actual public debt level in proportion to GDP, and a certain maximum sustainable public debt level for a given country, estimated on the past history of fiscal behavior of that country [Ostry et al., 2010; Gosh et al., 2013].

It should be noted that the results obtained by Reinhard and Rogoff [2010] have been challenged by Herndon, Ash and Pollin [2013], who demonstrate that while there is indeed a negative correlation between the debt level and the growth rate, the correlation is weaker than estimated by Reinhard and Rogoff, and the critical level is subsequently higher than 90%. In this paper, we do not enter this discussion and choose the 90% threshold as a matter of convention. One possible implication of this assumption is that the low number of asymmetric shocks in some highly indebted EA member countries, such as Belgium and Italy, may be partly due to the fact that the 90% threshold was not critical, because these countries were still able to use shock-absorbing fiscal policies. This indirectly confirms the existence of “non-revealed” shocks in these countries.

Correlations estimated with “Measure 2” shocks are generally weaker than with “Measure 1” shocks. This may be partly due to the fact that large deviations from moving averages are by definition proportionately smaller than large deviations from an exponential trend.

References


European Commission (1990), One market, one money: an evaluation of the potential benefits and costs of forming an economic and monetary union, *European Economy*, No. 44, October.

European Commission (2015), *Completing Europe's economic and monetary union*, report by: Jean-Claude Juncker in close cooperation with Donald Tusk, Jeroen Dijsselbloem, Mario Draghi, and Martin Schulz, Brussels, 22 June.


IMF (2013), Towards a fiscal union for the euro area, IMF Staff Discussion Note SDN/13/09, *International Monetary Fund*, September.


Size, Value, and Momentum in Polish Equity Returns: Local or International Factors?

Abstract

This paper tests the performance of the Capital Asset Pricing Model (CAPM) and the Fama-French three-factor and Carhart four-factor models on the Polish market. We use stock level data from April 2001 to January 2014 and find strong evidence for value and momentum effects, but only weak evidence for size premium. We formed portfolios double-sorted on size and book-to-market ratios, as well as on size and momentum, and we explain their returns with the above-mentioned asset pricing models. The CAPM is rejected and the three-factor and four-factor models perform well for the size and B/M sorted portfolios, but fail to explain returns on the size and momentum sorted portfolios. With the exception of the momentum factor, local Polish factors are not correlated with their European and global counterparts, suggesting market segmentation. Finally, the international value, size and momentum factors perform poorly in explaining cross-sectional variation in stock returns on the Polish market.

Keywords: value effect, size effect, momentum effect, Fama-French three-factor model, Carhart four-factor model, Polish market, asset pricing, market segmentation

JEL: G11, G12, G14, G15
Introduction

The value, size and momentum effects are not only well documented in the international markets, but also commonly applied in portfolio management, performance evaluation, and asset pricing. Research on the value effect dates back to Basu [1975, 1977, 1983], who found that low P/E stocks perform better than high P/E stocks. The value effect is the tendency of value stocks (stocks with low prices relative to their fundamentals) to outperform growth stocks (stocks with high prices relative to their fundamentals). Formal statistical evidence of the value effect was presented by Stattman [1980] and Rosenberg et al. [1985]. They use the book-to-market ratio as a value indicator. Davis et al. [1994] confirm the value effect in U.S. stock markets. Chan et al. [1991] and Capaul et al. [1993] confirm the value effect in markets outside the United States. The value effect is also found in stock returns by Fama and French [1998, 2012], Rouwenhorst [1999], Chui et al. [2010], and Asness, Moskowitz, and Pedersen [2013]. The size effect stems from observing the low capitalization stocks outperforming the large capitalization stocks and probably was first documented by Banz [1981]. Reinganum [1981], Blume and Stambaugh [1983], and Brown et al. [1983] confirm the evidence of the size effect by using a broader sample and decile portfolios. The size effect was also detected in several international markets by a variety of researchers [Herrera, Lockwood, 1994; Heston et al., 1999; Rouwenhorst, 1999; Fama, French, 2008; Michou et al., 2010; Asness et al., 2015; Hearn, 2016; Clarke et al., 2017]. Finally, the momentum effect refers to the phenomenon that stocks that have performed well during the past year tend to outperform the market in the future. It was initially documented by Jegadeesh and Titman [1993], who focused on a short-term investment horizon ranging from 3 to 12 months. Evidence for the momentum effect in returns from stocks on the U.S. and other international markets is put forward by, among others, Asness [1994], Fama and French [1998, 2012], Jegadeesh and Titman [2001], Rouwenhorst [1999], Grinblatt and Moskowitz [2004], Chui, Wei, and Titman [2010], Hearn [2016], Asness, Moskowitz, and Pedersen [2013], and Asness, Frazzini, Israel, Moskowitz, and Pedersen [2015].

Based on the above-described evidence, Fama and French [1992, 1993a] proposed a three-factor model and Carhart [1997] introduced a four-factor one. Both models became standard and are commonly used in financial and investment applications in developed markets, for example in portfolio performance evaluations [Kosowski et al., 2006; Fama, French, 2010]. The Carhart and Fama-French models to a great extent replaced the earlier popular CAPM model [Sharpe, 1964, 1966; Lintner, 1965; Mossin, 1966]. However, both multifactor models are not yet commonly employed in Poland.

There are two main reasons for this. The first is the lack of necessary data. There is a growing body of financial literature on integrated international asset pricing, which generally indicates that using only local pricing factors is applicable in asset pricing, contrary to regional or global asset pricing factors [Griffin, 2002; Hou et al., 2011; Fama,
French, 2012; Cakici et al., 2013]. Unfortunately, the high quality and comprehensive data on the Fama-French and Carhart factors for the Polish market are not easily accessible. The second is the lack of developed empirical literature on cross-sectional pricing in the Polish market. There is limited research on asset pricing anomalies, which generally confirms value [Borys, Zemcik, 2009; Lischewski, Voronkova, 2012; Zaremba, Konieczka, 2014, 2015a, 2015b], size [Borys, Zemcik, 2009; Welc, 2012; Lischewski, Voronkova, 2012; Sekuła, 2013; Zaremba, Konieczka, 2014, 2015a, 2015b], and momentum effects [Szyszka, 2006; Żebrowska-Suchodolska, Witkowska, 2008; Zaremba, Konieczka, 2014, 2015a, 2015b]. Some literature attempts to apply the Fama-French model [Czapkiewicz, Skalna, 2010; Olbryś, 2010; Urbański, 2012; Waszczuk, 2013a, 2013b]. However, we are not aware of any study that comprehensively examines applicability of the Fama-French and Carhart models based on local and international factors on the Polish market.

The crucial purpose of this study is to at least partially fill the above-described knowledge gap. In other words, this paper aims to comprehensively explore and compare the applicability of three pricing models in the Polish market: the CAPM, the Fama-French three-factor model, and the Carhart four-factor model. It contributes to the economic literature in four ways. First, it investigates the value, size, and momentum premium in the Polish market. Second, it examines correlations between Polish and international asset pricing factors. Third, it tests the explanatory power of the three above-mentioned asset pricing models for Poland. Finally, it compares the performance of multifactor models based on local and international factors.

The principal findings can be summarized as follows. First, we find positive and significant value and momentum premium, but positive and non-significant size premium. Second, local Polish pricing factors are not correlated with their European or global counterparts, with the exception of momentum. Third, the Fama-French and Carhart models explain the B/M and size double-sorted portfolios, but fail to explain the portfolios double-sorted on momentum and size. Fourth, the multifactor models based on international factors perform poorly in the Polish market.

We use stock level data from Poland. The sample period is 2001–2014. In the asset pricing tests, we explain the returns of size and B/M and size and momentum double-sorted portfolios typical in the literature. We evaluate their performance using three well-known models.

The structure of the paper is as follows. In section 2 we discuss the examined asset-pricing models and the statistical tests we employ to examine the model performance. Next, in section 3, we analyze the data and the variables. Our findings are presented in section 4, and section 5 concludes this paper.
Research Methods and Asset Pricing Models

We test the explanatory power of three distinct pricing models, which are all estimated using cross-sectional data. The first model is the classical Capital Asset Pricing Model [Sharpe, 1964, 1966; Lintner, 1965; Mossin, 1966]. The model assumes that asset returns depend only on the market portfolio and is described by the following regression equation:

\[ R_{i,t} = \alpha_i + R_{f,t} + \beta_{rm,i} \cdot (R_{mt} - R_{f,t}) + \epsilon_{i,t}, \] (1)

where \( R_{i,t}, R_{m,t}, \) and \( R_{f,t} \) are returns on the analyzed asset \( i \), market portfolio and risk-free returns at time \( t \), and \( \alpha_i \) and \( \beta_{rm,i} \) are regression parameters. The \( \alpha_i \) intercept measures the average abnormal return (the so-called Jensen-alpha). The second model is the Fama-French three-factor model [Fama, French, 1992, 1993a]:

\[ R_{i,t} = \alpha_i + R_{f,t} + \beta_{rm,i} \cdot (R_{m,t} - R_{f,t}) + \beta_{SMB,i} \cdot SMB_t + \beta_{HML,i} \cdot HML_t + \epsilon_{i,t}, \] (2)

where \( \beta_{rm,i} \) and \( \beta_{SMB,i}, \beta_{HML,i} \) are analogical to the CAPM beta, but it is not equal to it. The \( \beta_{SMB,i}, \beta_{HML,i} \) are exposed to SMB \( i \) and HML \( i \) risk factors, which denote returns from zero-cost arbitrage portfolios. SMB \( i \) is the difference in returns on diversified portfolios of small and large caps at time \( t \), whereas HML \( i \) is, in general, the difference in returns on portfolios of diversified value (high B/V) and growth (low B/V) stocks. In other words, SMB and HML are returns on zero-cost market-neutral long/short portfolios that are formed based on size and value characteristics. The Fama-French model was tested multiple times, in particular with respect to the U.S. market [Fama, French, 1996; Daniel, Titman, 1997; Davis et al., 2000]. For the most part, the model was also tested on developed markets and the global market. At the same time, tests on emerging markets were fairly infrequent. With respect to Poland, research on the applicability of the model to the Warsaw Stock Exchange was conducted by Kowerski [2008] and Czapkiewicz and Skalna [2010].

The last model is the four-factor model, which was originally introduced by Carhart [1997]. Its corresponding regression equation is:

\[ R_{i,t} = \alpha_i + R_{f,t} + \beta_{rm,i} \cdot (R_{m,t} - R_{f,t}) + \beta_{SMB,i} \cdot SMB_t + \beta_{HML,i} \cdot HML_t + \beta_{WML,i} \cdot WML_t + \epsilon_{i,t}, \] (3)

The model additionally incorporates momentum returns measured by returns on so-called winner and loser portfolios, which were used in the initial studies of this anomaly [Jegadeesh, Titman, 1993]. The WML \( t \) denotes the difference between returns on diversified winner and loser portfolios during the previous year. The model was developed by Carhart
and was later tested by Jegadeesh [2000], Liew and Vassalou [2000], Kim and Kim [2003], L’Her, Masmoudi, and Suret [2004], Bello [2007], and Lam, Li, and So [2009].

All the regression models are estimated using OLS regression and tested in a parametric way. Following the extensive literature on the subject, we test the models by assessing the performance of various value, momentum, and size sorted portfolios, which are described in detail in the data section. We examine whether the model application to a certain portfolio leaves a statistically significant intercept unexplained. In order to find whether the intercepts in a group of portfolios are statistically different from zero, we evaluate the models’ performance with the popular GRS test statistic, as suggested by Gibbons, Ross, and Shanken [1989]. The test statistic is given by:

\[
GRS = \frac{T}{N} \left( \frac{T-N-L}{T-L-1} \right) \hat{\alpha}' \hat{\Sigma}^{-1} \hat{\alpha} \cdot \left[ 1 + E_T(f) \hat{\Omega}^{-1} E_T(f) \right]^{-1} \sim F_{N,T-N-L}, \tag{4}
\]

where \( T \) is the length of the time-series (sample size), \( N \) is the number of portfolios to be explained in the examined group, and \( L \) denotes the number of explanatory factors. \( E_T(f) \) is a vector of expected returns to asset pricing factors (estimated as a simple average during the investigated period; see Cochrane [2005, p. 231]), \( \hat{\Omega} \) is a covariance matrix of the asset pricing factors, \( \hat{\alpha} \) is a vector of regression intercepts, and \( \hat{\Sigma} \) is a residual covariance matrix in the sample. The test’s critical values are obtained from Fisher’s distribution with \( N \) and \( T-N-L \) degrees of freedom.

### Data Sources and Preparation

Bloomberg was the primary data source and stock level data on all the companies in the Polish market available in Bloomberg were used. Both listed and delisted companies were analyzed so as to avoid survivorship bias. The primary sample period comprised April 2001 to January 2014. To include a company into the sample in a given time, all necessary characteristics to compute a pricing factor had to be obtained (for example: size and B/M for the HML factor). The number of companies in the sample grew from 119 to 827 and the average number was 308. Earlier data were not used due to the fact that pool of small companies did not allow reasonable portfolios to be formed. Moreover, in some cases, we had to curtail the data period from November 2011 to January 2014, as there were too few companies available in earlier months to form the necessary number of portfolios. These cases are precisely indicated in the results description.

According to the models presented in equations (1), (2), and (3), we used four distinct pricing factors: \( R_m-R_f \), HML, SMB, and WML. Whenever we refer to the European or
global pricing factors, data from Kenneth French’s website were used; however, factors from the Polish market are computed for the purpose of this paper. The $R_m - R_f$ is the difference between the return on the WIG Index (the broadest Polish equity market total return index, which encompasses almost the entire Polish market) and the 1-month Warsaw Interbank Bid Rate (WIBID). Furthermore, all the excess returns in the study are calculated during the 1-month WIBID rate.

In order to calculate the remaining factors (HML, SMB, and WML), we sorted all the stocks each month according to three distinct characteristics:

– B/M ratio: the book value of equity to market value of equity,
– size: the total stock market capitalization of the equity, and
– momentum: the cumulative stock return from time $t-12$ to $t-25$.

The computational methodology of the traits described above is consistent with the methodology in similar asset pricing studies [Fama, French, 2012; Cakici et al., 2013; De Groot et al., 2012].

First, for presentational purposes, returns were calculated on B/M, size, and momentum sorted quantile portfolios. Stocks were divided into five independent quantiles based on B/M, size, and momentum. For each month, we calculated the 20, 40, 60, and 80 percentiles for B/M, size, and momentum breakpoints. Based on these breakpoints, five distinct quantile portfolios were created in the case of each characteristic. The value-weighting scheme was used.

Next, the precise returns on HML, SMB, and WML factors were calculated. Again, the computational methodology is consistent with most of the popular asset pricing studies. Initially, the stocks were divided into two size portfolios based on their stock capitalization. We define the size breakpoint as the median size of all the stocks in the Polish market in a given month. The stocks above the median were classified as the large stocks, and the remaining ones as the small stocks. In other words, the number of stocks in both portfolios is usually equal. It is important to note that in the Polish market there is a considerable and increasing number of very-low capitalization stocks, so the above-described large-cap portfolio constituted almost 97% of the stock market capitalization in the beginning of the study and more than 99% in the end. Again, for all stocks, we determined the standard top 30% (value), middle 40% (neutral), and bottom 30% (value) breakpoints. In other words, stocks with the highest B/M ratios are regarded as value stocks and stocks with the lowest B/M ratios as growth stocks. The computed B/M breakpoints were applied to the big and small stocks, so six groups of stocks that emerged from the double-sorts on size and B/M were developed. Next, based on the described division, we formed six value-weighted portfolios, which were denoted by BV, BN, BG, SV, SN, and SG, where B and S refer to big or small, and V, N and G refer to value, neutral, and growth. The HML return in month $t$ was estimated as the difference between the equal weighted average returns of small and large value stocks ($R_{SV}$, $R_{BV}$) and the equal weighted average returns of small and big growth stocks ($R_{SG}$, $R_{BG}$):
Additionally, we computed HML factors specific for small stocks ($HML_s$) and big stocks ($HML_b$), which are basically computed within the small and big subgroups:

$$HML_{s,t} = R_{SV,t} - R_{SG,t}, \quad (6)$$

$$HML_{b,t} = R_{BV,t} - R_{BG,t} \quad (7)$$

The next factor, $SMB$, is the difference between the equal weighted average return of the three small-cap portfolios and the equal weighted average return of the three large-cap portfolios:

$$SMB_t = \frac{1}{3} \left( R_{SV,t} + R_{SN,t} + R_{SG,t} \right) - \frac{1}{3} \left( R_{BV,t} + R_{BN,t} + R_{BG,t} \right). \quad (8)$$

The calculation of the $WML$ factor is almost identical to the $HML$ factor, but instead of B/M ratio momentum was used. We determined the standard top 30% (winners), middle 40% (neutral), and bottom 30% (losers) breakpoints. The momentum breakpoints were applied to the big and small stocks, so we created six groups of stocks that emerge from double-sorts on size and momentum. Next, we built six value-weighted portfolios, which are denoted by BW, BN, BL, SW, SN, and SL, where B and S refer to big or small, and W, N, and L refer to winners, neutral, and losers. The $WML$ return in month $t$ was computed as the difference between the equal weighted average returns of small and large winner stocks ($RSW, RBW$) and the equal weighted average returns of small and big loser stocks ($RSL, RBL$):

$$WML_t = \frac{1}{2} \left( R_{SW,t} + R_{BW,t} \right) - \frac{1}{2} \left( R_{SL,t} + R_{BL,t} \right) \quad (9)$$

Furthermore, we computed the $WML$ factors specific for small stocks ($WML_s$) and big stocks ($WML_b$), which were basically computed within the small and big subgroups:

$$WML_{s,t} = R_{SW,t} - R_{SL,t}, \quad (10)$$

$$WML_{b,t} = R_{BW,t} - R_{BL,t} \quad (11)$$

The models described in equations (1), (2), and (3) were tested against two distinct groups of portfolios: 25 sorts on B/M and size and 25 sorts on momentum and size. Formation of those $5 \times 5$ size-momentum and size-B/M portfolios is analogical to the formation of $2 \times 3$ size-momentum and size-B/M portfolios necessary to obtain the asset pricing factors. Beginning with the B/M and size portfolio, we initially sorted all stocks on
the book-to-market ratio to find the 20, 40, 60, and 80 percentile B/M breakpoints. Next, the procedure was repeated for the size factor, so that companies were sorted according to their stock market capitalization and 20, 40, 60, and 80 percentile size breakpoints were found. Finally, intersecting the two independent sorts on the B/M and size, all stocks were placed into one of the value-weighted B/M-size portfolios. Formation of the 25 momentum-size portfolios was identical, however, with one obvious exception: the momentum ratio (cumulative return in months t-12 to t-2) was used, instead of the B/M one.

**Results and Interpretation**

This section, summarizes statistics for value, size, and momentum sorted portfolios and for asset pricing factors in the Polish market. Next, it turns to the 25 portfolios formed on B/M, size and momentum. Then, it investigates the interdependencies between Polish and international pricing factors. Finally, it reports on the results of asset pricing tests, which make use of asset pricing factors such as the RHS explanatory returns and the 25 double-sorted portfolios such as the LHS assets in the regressions.

**TABLE 1. Excess returns on quantile portfolios sorted on B/M, size, and momentum**

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Size</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.96</td>
<td>0.76</td>
<td>0.37</td>
<td>0.43</td>
<td>0.22</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>9.56</td>
<td>8.52</td>
<td>7.63</td>
<td>7.62</td>
<td>6.72</td>
</tr>
<tr>
<td><strong>B/M ratio</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.14</td>
<td>0.08</td>
<td>0.54</td>
<td>0.27</td>
<td>0.95</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>7.30</td>
<td>7.71</td>
<td>6.68</td>
<td>6.75</td>
<td>8.67</td>
</tr>
<tr>
<td><strong>Momentum</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-1.21</td>
<td>-0.55</td>
<td>-0.21</td>
<td>0.36</td>
<td>1.31</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>10.11</td>
<td>9.18</td>
<td>8.09</td>
<td>7.01</td>
<td>7.86</td>
</tr>
</tbody>
</table>

The table reports the means and standard deviations of excess log-returns on quantile portfolios formed on book equity to market equity (B/M) ratios, size (market capitalization), and momentum. For portfolios created at the end of month t, the lagged momentum return is the stock cumulative return from time t-12 to t-2. Computations were based on monthly time-series. All firms were sorted into 10 size groups and value weighted to obtain quantile portfolios. All returns were calculated using stock level data from Bloomberg. The data period is 04/30/2001–01/31/2014.

Source: own study.

Table 1 reports excess returns for factor sorted portfolios. Starting with the B/M ratio, during the entire research period high B/M stocks had larger returns than the low B/M stocks. The top B/M portfolios mean excess log-return was 0.95% (value weighted), whereas
the low B/M means were 0.14%. What is interesting is that the extreme B/M portfolios seem to be slightly riskier in terms of standard deviation than the neutral B/M portfolios. Although value stocks appear to perform better than growth stocks, the exact size of this dominance seems to be highly time-variant.

The superior performance of small stocks is also evident on the Polish market. The smallest stocks’ mean excess log-return was 1.96% and the same parameter for large-caps was 0.22%. Nonetheless, small-caps also seemed to be riskier with a standard deviation of 9.56%, contrary to a standard deviation of 6.72% for the biggest companies.

The momentum effect seems to be the strongest of the three examined anomalies. In the case of value weighted portfolios, the top-momentum portfolios earned a mean monthly excess log-return of 1.31%, whereas the bottom momentum portfolio on average lost 1.21%. What is quite interesting is that the better performing winners were actually less risky. For example, turning to the value weighted portfolios, the winners’ excess log-returns’ standard deviation was 7.86%, whereas for the losers it amounts to 10.11%.

The value effect was particularly strong among the small companies (Table 2). This observation is consistent with previous studies, which show that the value premium is strongest in the small-cap universe [Kothari et al., 1995; Loughran, 1997; Dhatt et al., 1999; Fama, French, 2006; Asness et al., 2015]. For the value weighted portfolios, the high B/M small-caps earned 1.34% excess log-returns monthly, whereas the low-B/M small-caps earned 0.24%. In the case of the large-caps, this value amounted to 0.69% and 0.03%, so the difference was much smaller. The domination of value premium across the small firms was more or less time-invariant.

| TABLE 2. Excess returns on portfolios from 2 × 3 sorts on B/M, momentum, and size |
|-----------------------------------------------|---------------|---------------|---------------|---------------|---------------|---------------|
|                                               | Mean           | Standard deviation |
|                                               | Low    | Medium  | High   | Low    | Medium  | High   |
| B/M ratio                                     |         |         |        |        |         |        |
| Small                                         | 0.24   | 0.72    | 1.34   | 9.71   | 8.01    | 8.71   |
| Big                                           | 0.03   | 0.28    | 0.69   | 7.02   | 6.76    | 7.81   |
| Momentum                                      |         |         |        |        |         |        |
| Small                                         | 0.00   | 0.71    | 1.39   | 9.51   | 7.88    | 9.23   |
| Big                                           | −1.11  | −0.33   | 1.05   | 9.43   | 7.7     | 7.26   |

The table reports the means and standard deviations of excess log-returns on six portfolios formed on the book equity to market equity (B/M) ratios, momentum (total return in months t-12 to t-2), and size (market capitalization). All firms were sorted into two size groups and three momentum groups. We intersected the two sorts on size and three on value and value weight to obtain six portfolios. The computations were based on monthly time-series. All returns were calculated using stock level data from Bloomberg. The data period is 04/30/2001–01/31/2014. Source: own study.
Surprisingly the situation is quite different in the case of momentum effect. Recent research shows that momentum effect is generally stronger for small companies [Hong et al., 2000; Fama, French, 2012]. By contrast, the obtained results show that in Poland the differences in mean excess returns was larger in the large-cap universe. For value weighted portfolios, differences between winners and losers equaled 1.39% in the small-cap universe and 2.16% in the large-cap universe. These differences were naturally time-variant; however, in both investigated periods the momentum effect was stronger among larger companies.

TABLE 3. Excess returns on portfolios from 5 × 5 sorts on B/M ratio, momentum, and size

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th></th>
<th>Standard deviation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>B/M ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>1.40</td>
<td>1.15</td>
<td>1.47</td>
<td>2.00</td>
</tr>
<tr>
<td>2</td>
<td>0.20</td>
<td>0.67</td>
<td>1.42</td>
<td>1.07</td>
</tr>
<tr>
<td>3</td>
<td>0.60</td>
<td>0.52</td>
<td>0.68</td>
<td>0.25</td>
</tr>
<tr>
<td>4</td>
<td>0.74</td>
<td>−0.03</td>
<td>0.36</td>
<td>0.28</td>
</tr>
<tr>
<td>Big</td>
<td>−0.21</td>
<td>0.06</td>
<td>0.28</td>
<td>0.83</td>
</tr>
</tbody>
</table>

|            | Low  | 2        | 3        | 4        | High |
| Momentum   |      |          |          |          |      |
| Small      | 2.57 | 3.00     | 1.86     | 2.07     | 1.77 |
| 2          | 0.19 | 1.57     | 1.01     | 0.23     | 1.76 |
| 3          | −1.23| −0.89    | 0.28     | 1.49     | 1.95 |
| 4          | −0.98| −0.26    | 0.45     | 0.67     | 2.25 |
| Big        | −1.08| −0.34    | 0.17     | −0.26    | 1.44 |

The table reports the means and standard deviations of excess returns on 25 portfolios formed on the book value to market value ratio (B/M), momentum (cumulative return in t-12 to t-2), and size (market capitalization). All the firms were sorted into five size groups and five B/M groups. We intersected the five sorts on size and B/M and value weight in order to obtain 25 portfolios. The computations were based on monthly time-series. All the returns were calculated using stock level data from Bloomberg. The data period is 11/30/2002–01/31/2014.

Source: own study.

The results of the 5 × 5 double sorted B/M-size and momentum-size (Table 3) portfolios generally echo the numbers in Tables 1 and 2. Concentrating on the 25 value-weighted portfolios from 5 × 5 sorts on B/M and size, one can see that the high B/M small firms performed particularly well compared to the remaining portfolios and the mean excess log-returns is equal to 3.29%. On the other hand, the performance of large-cap low B/M stocks is particularly poor and the excess returns amounted to only −0.21%. Additionally, it is important to note that the distribution of the means was quite uneven. For example, although the small high B/M stocks performed better than the large high B/M stocks,
both portfolios outperformed mid-size high B/M stocks. The pattern is similar in the small-cap universe. The small-cap stocks with neutral B/M tended to have lower returns than both stocks with high and low B/M. The results for the equal weighted B/M-size portfolios generally resembled the above-described statistics.

The 25 portfolios from sorts on size and momentum confirm the confusing anomaly with large-cap momentum premium magnitude. Among the smallest companies (top row of the matrices), it is the low momentum quantile which performed better. In the universe of large companies, the situation seems to be rather in line with previous studies on the momentum, and the winner stocks performed better than the losers.

**TABLE 4. Means, standard deviations, and t-statistics for asset pricing factors**

<table>
<thead>
<tr>
<th></th>
<th>Rm-Rf</th>
<th>SMB</th>
<th>HML</th>
<th>WML</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.56</td>
<td>0.43</td>
<td>0.71</td>
<td>1.39</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>6.60</td>
<td>5.17</td>
<td>5.39</td>
<td>6.56</td>
</tr>
<tr>
<td>t-stat</td>
<td>1.05</td>
<td>1.03</td>
<td>1.66</td>
<td>2.64</td>
</tr>
</tbody>
</table>

The table describes the means and standard deviations of asset pricing factors in Poland. Rm-Rf is the return on the WIG Index minus 1-month WIBID rate. SMB is the small minus big factor, HML is the high minus low factor, and WML is the momentum factor. S and B subscripts stand for small stocks and big stocks, respectively. The computations are based on monthly time-series. All the returns were calculated using stock level data from Bloomberg. The data period is 04/30/2001–01/31/2014. The table also reports the t-statistics (t-stat).

Source: own study.

Table 4 shows means and standard deviation for factor returns on the Polish market. The $R_m-R_f$ is equal to 0.56%, and is not economically significant and highly time-variant. Nonetheless, the lack of statistical significance may be a result of the relatively short investigation period. Turning to the SMB factor, it was equal to 0.43%. Its size is relatively small compared to WML and HML factors and statistically insignificant. These observations echo the study of Fama and French [2012], who also did not find evidence for the SMB premium in the post-1990 period.

Evidence for the HML factor seems to be stronger. This premium was positive and statistically significant at the 90% level. Its monthly effect was 0.71%. The WML factor appeared to be the strongest of all. With the test statistic of 2.64%, it was significant at the 99% level. Its monthly effect was positive in all subperiods and equal to 1.39% in the entire time-series sample.

We also tested the correlation between asset pricing factors in Poland and the international markets, which should indicate whether the use of international pricing factors in the Polish market is appropriate or not. On the other hand, they are important for investors pursuing size, value, and momentum strategies with a geographical focus.
TABLE 5. **Correlation coefficients and t-statistics for global and local asset pricing factors**

### Panel A: Poland vs Poland

<table>
<thead>
<tr>
<th>Correlation coefficients</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mkt-RF&lt;sub&gt;PL&lt;/sub&gt;</td>
<td>1.00</td>
</tr>
<tr>
<td>SMB&lt;sub&gt;PL&lt;/sub&gt;</td>
<td>1.00</td>
</tr>
<tr>
<td>HML&lt;sub&gt;PL&lt;/sub&gt;</td>
<td>1.00</td>
</tr>
<tr>
<td>WML&lt;sub&gt;PL&lt;/sub&gt;</td>
<td>1.00</td>
</tr>
</tbody>
</table>

| Mkt-RF<sub>PL</sub> | - | -0.19 | 1.92 | -2.55 |
| SMB<sub>PL</sub> | - | -2.95 | 0.52 |
| HML<sub>PL</sub> | - | -5.87 |
| WML<sub>PL</sub> | - |

### Panel B: Poland vs Europe

<table>
<thead>
<tr>
<th>Correlation coefficients</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mkt-RF&lt;sub&gt;EU&lt;/sub&gt;</td>
<td>0.67</td>
</tr>
<tr>
<td>SMB&lt;sub&gt;EU&lt;/sub&gt;</td>
<td>0.20</td>
</tr>
<tr>
<td>HML&lt;sub&gt;EU&lt;/sub&gt;</td>
<td>0.18</td>
</tr>
<tr>
<td>WML&lt;sub&gt;EU&lt;/sub&gt;</td>
<td>0.46</td>
</tr>
</tbody>
</table>

| Mkt-RF<sub>PL</sub> | 11.24 | 1.00 | 4.19 | -5.76 |
| SMB<sub>PL</sub> | 2.52 | 0.66 | -0.59 |
| HML<sub>PL</sub> | 2.24 | -3.65 |
| WML<sub>PL</sub> | 6.32 |

### Panel C: Poland vs world

<table>
<thead>
<tr>
<th>Correlation coefficients</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mkt-RF&lt;sub&gt;GL&lt;/sub&gt;</td>
<td>0.71</td>
</tr>
<tr>
<td>SMB&lt;sub&gt;GL&lt;/sub&gt;</td>
<td>0.20</td>
</tr>
<tr>
<td>HML&lt;sub&gt;GL&lt;/sub&gt;</td>
<td>0.09</td>
</tr>
<tr>
<td>WML&lt;sub&gt;GL&lt;/sub&gt;</td>
<td>0.48</td>
</tr>
</tbody>
</table>

| Mkt-RF<sub>PL</sub> | 12.30 | 2.01 | 1.97 | -4.93 |
| SMB<sub>PL</sub> | 2.58 | 0.10 | -0.51 |
| HML<sub>PL</sub> | 1.14 | -4.02 |
| WML<sub>PL</sub> | 6.79 |

The table presents the correlation coefficient between pricing factors in the Polish, European, and global settings. MKT-RF is the return on the WIG Index minus 1-month WIBID rate. SMB is the small minus big factor, HML is the high minus low factor, and WML is the momentum factor. S and B subscripts stand for small stocks and big stocks, respectively. The computations are based on monthly time-series. All the returns were calculated using stock level data from Bloomberg. The data period is 04/30/2001–12/31/2013. The table also reports on the t-statistics (t-stat). The PL subscript refers to the Polish factors, whereas the EU and GL refer to the European and global pricing factors, respectively. The correlation coefficients of Polish asset pricing factors in vertical columns were computed against the Polish (panel A), European (panel B), and global (panel C) factors. Factors for Europe and the world were obtained from Kenneth French’s website.

**Source:** own study.

First, we focus on intra-country correlations between pricing factors (Table 5). The correlations were generally low and ranged from −0.23 to 0.15. However, an interesting exception is the correlation between value and momentum (HML and WML). It was negative, amounting to −0.43, and its corresponding test statistic was −5.87. This observation is consistent with the studies of Asness, Moskowitz, and Pedersen [2013], and Cakici, Fabozzi,
and Tan [2013], who find that value and momentum factors are negatively correlated, making it possible to combine them into efficient portfolios.

### Table 6. Intercepts from value-pricing models to explain monthly excess returns on portfolios from 5 x 5 sorts on size and B/M

#### Panel A: CAPM

<table>
<thead>
<tr>
<th>Intercept</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>2</td>
</tr>
<tr>
<td>Small</td>
<td>2.27</td>
</tr>
<tr>
<td>2</td>
<td>0.20</td>
</tr>
<tr>
<td>3</td>
<td>0.43</td>
</tr>
<tr>
<td>4</td>
<td>0.35</td>
</tr>
<tr>
<td>Big</td>
<td>-0.60</td>
</tr>
</tbody>
</table>

#### Panel B: three-factor model

<table>
<thead>
<tr>
<th>Intercept</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>2</td>
</tr>
<tr>
<td>Small</td>
<td>2.13</td>
</tr>
<tr>
<td>2</td>
<td>-0.09</td>
</tr>
<tr>
<td>3</td>
<td>0.12</td>
</tr>
<tr>
<td>4</td>
<td>0.09</td>
</tr>
<tr>
<td>Big</td>
<td>-0.29</td>
</tr>
</tbody>
</table>

#### Panel C: four-factor model

<table>
<thead>
<tr>
<th>Intercept</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>2</td>
</tr>
<tr>
<td>Small</td>
<td>2.41</td>
</tr>
<tr>
<td>2</td>
<td>-0.17</td>
</tr>
<tr>
<td>3</td>
<td>0.02</td>
</tr>
<tr>
<td>4</td>
<td>-0.18</td>
</tr>
<tr>
<td>Big</td>
<td>-0.42</td>
</tr>
</tbody>
</table>

The table describes the intercepts and t-statistics of the intercepts of 25 portfolios formed on the book value to market value ratio (B/M) and size (market capitalization). All firms were sorted into five size groups and five B/M groups. We intersected the five sorts on size and B/M and value weight to obtain 25 portfolios. The regressions used the CAPM (panel A), three-factor (panel B), and four-factor (panel C) models with Polish factors to explain excess returns on 25 double-sorted portfolios formed from independent size and B/M sorts. The calculations were based on monthly time-series. All the returns were determined using stock level data from Bloomberg. The data period is 11/30/2002–12/31/2013.

Source: own study.
Second, referring to correlations with the foreign asset pricing factors, we find that the correlation coefficients between Polish and foreign HML and SMB factors were generally very low, but positive and ranged from 0.01 to 0.20. Again, the behavior of the WML factor appears to be quite different. On the one hand, it was positively correlated at the 0.46–0.48 level with the European and global pricing factors. On the other hand, it was negatively correlated with the international $R_m-R_f$ and HML factors.

Table 6 reports on intercepts and their test statistics of 25 sorts on B/M and size tested against the Polish factors-based CAPM, three-, and four-factor models. Generally, the CAPM does the poorest job in explaining abnormal returns. The model leaves unexplained intercepts for a few small-cap (the top row of the matrix) and high B/M (the right column of the matrix) portfolios. This failure to explain returns is almost entirely corrected by the three- and four-factor models. Both models leave only two portfolios with statistically significant positive intercepts: the two top B/M portfolios in the smallest companies’ quantile universe. The alphas of the most extreme high B/M small-cap portfolios were 1.90% (three-factor) and 2.41% (four-factor) with corresponding t-statistics of 2.60 and 3.16.

Explaining returns of the 25 size and momentum sorted portfolios is difficult within the investigated models (Table 7). The CAPM and three-factor models are unable to explain returns of numerous small-cap-, loser-, and winner-portfolios. The matrices in panels A and B reveal many statistically significant alphas. This improves after applying the four-factor model, which included the momentum factor. Most portfolios show no notable intercept patterns. However, the model fails to explain the reversed momentum spread among small companies and significant negative returns to loser portfolios. Apart from the anomalous low-momentum mid-cap portfolio, this leaves a few portfolios with intercepts statistically different than 0 at the 95% level, which ranged to 1.82%.

Panel A of Table 8 summarizes the research on whether the Polish, European and global asset pricing factors explain returns on the B/M and size double-sorted portfolios. First, the performance of the foreign factors was rather weak. The GRS test statistics reject all the models at the 95% confidence level. The models explain only 22–30% of variation in the portfolios’ excess returns and the average absolute intercept varies from 0.62 to 0.70. Turning to the local factors, the Polish CAPM model is also rejected, with the GRS statistic of 8.27. However, the model explains 46% of the variation, which is much more than its European and global counterparts. The corresponding average absolute intercept was 0.73%. When the two additional factors — HML and SMB — were applied, the performance of the model improved significantly. The average absolute intercept fell to 0.53 and the $R^2$ rose to 60%. The corresponding GRS statistics was 1.14, which means that the model is not rejected. Furthermore, inclusion of the momentum factor to the model did not introduce much improvement. The $R^2$ increased by only 0.6 percentage points and the average absolute and GRS statistic actually marginally rose. Nonetheless, the model is not rejected based on the GRS statistic.
TABLE 7. Intercepts from value-pricing models to explain monthly excess returns on portfolios from $5 \times 5$ sorts on size and momentum

Panel A: CAPM

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Intercept</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>High</td>
<td>Low</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>High</td>
</tr>
<tr>
<td>Small</td>
<td>2.45</td>
<td>3.44</td>
<td>1.57</td>
<td>1.89</td>
<td>2.80</td>
<td>2.92</td>
<td>2.84</td>
<td>1.65</td>
<td>1.76</td>
<td>2.39</td>
</tr>
<tr>
<td>2</td>
<td>-0.35</td>
<td>1.54</td>
<td>0.55</td>
<td>0.57</td>
<td>0.81</td>
<td>-0.52</td>
<td>1.85</td>
<td>0.84</td>
<td>0.86</td>
<td>0.84</td>
</tr>
<tr>
<td>3</td>
<td>-1.67</td>
<td>-0.75</td>
<td>-0.02</td>
<td>1.41</td>
<td>1.92</td>
<td>-2.80</td>
<td>-1.15</td>
<td>-0.05</td>
<td>2.13</td>
<td>2.43</td>
</tr>
<tr>
<td>4</td>
<td>-1.33</td>
<td>-0.74</td>
<td>0.15</td>
<td>1.45</td>
<td>2.29</td>
<td>-2.07</td>
<td>-1.38</td>
<td>0.30</td>
<td>2.09</td>
<td>2.89</td>
</tr>
<tr>
<td>Big</td>
<td>-1.12</td>
<td>-0.49</td>
<td>-0.57</td>
<td>-0.34</td>
<td>0.93</td>
<td>-1.53</td>
<td>-0.80</td>
<td>-1.42</td>
<td>-0.84</td>
<td>1.89</td>
</tr>
</tbody>
</table>

Panel B: three-factor model

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Intercept</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>High</td>
<td>Low</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>High</td>
</tr>
<tr>
<td>Small</td>
<td>1.47</td>
<td>2.16</td>
<td>0.69</td>
<td>1.10</td>
<td>1.96</td>
<td>2.00</td>
<td>2.03</td>
<td>0.85</td>
<td>1.13</td>
<td>1.88</td>
</tr>
<tr>
<td>2</td>
<td>-1.50</td>
<td>0.37</td>
<td>-0.21</td>
<td>-0.21</td>
<td>0.22</td>
<td>-2.75</td>
<td>0.54</td>
<td>-0.40</td>
<td>-0.37</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>-2.39</td>
<td>-1.65</td>
<td>-0.62</td>
<td>0.74</td>
<td>1.30</td>
<td>-4.37</td>
<td>-2.86</td>
<td>-1.39</td>
<td>1.27</td>
<td>2.27</td>
</tr>
<tr>
<td>4</td>
<td>-2.27</td>
<td>-1.39</td>
<td>-0.47</td>
<td>1.12</td>
<td>1.71</td>
<td>-3.98</td>
<td>-2.83</td>
<td>-1.12</td>
<td>1.68</td>
<td>2.35</td>
</tr>
<tr>
<td>Big</td>
<td>-1.66</td>
<td>-1.13</td>
<td>-0.58</td>
<td>-0.20</td>
<td>1.09</td>
<td>-2.29</td>
<td>-1.93</td>
<td>-1.42</td>
<td>-0.47</td>
<td>2.27</td>
</tr>
</tbody>
</table>

Panel C: four-factor model

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Intercept</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>High</td>
<td>Low</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>High</td>
</tr>
<tr>
<td>Small</td>
<td>2.07</td>
<td>2.09</td>
<td>0.41</td>
<td>0.62</td>
<td>1.27</td>
<td>2.72</td>
<td>1.86</td>
<td>0.48</td>
<td>0.61</td>
<td>1.17</td>
</tr>
<tr>
<td>2</td>
<td>-1.12</td>
<td>0.97</td>
<td>-0.56</td>
<td>-1.02</td>
<td>-1.34</td>
<td>-1.97</td>
<td>1.37</td>
<td>-1.01</td>
<td>-1.83</td>
<td>-1.60</td>
</tr>
<tr>
<td>3</td>
<td>-1.82</td>
<td>-1.19</td>
<td>-0.19</td>
<td>0.60</td>
<td>0.26</td>
<td>-3.26</td>
<td>-2.00</td>
<td>-0.42</td>
<td>0.98</td>
<td>0.48</td>
</tr>
<tr>
<td>4</td>
<td>-1.37</td>
<td>-0.92</td>
<td>-0.19</td>
<td>0.95</td>
<td>0.68</td>
<td>-2.48</td>
<td>-1.82</td>
<td>-0.42</td>
<td>1.33</td>
<td>0.95</td>
</tr>
<tr>
<td>Big</td>
<td>-0.56</td>
<td>-0.15</td>
<td>-0.54</td>
<td>-0.65</td>
<td>0.16</td>
<td>-0.79</td>
<td>-0.26</td>
<td>-1.23</td>
<td>-1.52</td>
<td>0.37</td>
</tr>
</tbody>
</table>

The table describes the intercepts and t-statistics of the intercepts of 25 portfolios formed on momentum and size (market capitalization). For portfolios created at the end of month $t$, the lagged momentum return is the stock cumulative return from time $t-12$ to $t-2$. All the firms were sorted into five size groups and five momentum groups. We intersected the five sorts on size and momentum and value weight to obtain 25 portfolios. The regressions used the CAPM (panel A), three-factor (panel B), and four-factor (panel C) models with Polish factors to explain the excess returns on 25 double-sorted portfolios formed from independent size and value sorts. The computations were based on monthly time-series. All the returns were calculated using stock level data from Bloomberg. The data period is 11/30/2002–12/31/2013.

Source: own study.
TABLE 8. **Summary to explain monthly excess returns on portfolios from 5 × 5 sorts on size and value and 5 × 5 sorts on size and momentum**

Panel A: size and B/M

|                | GRS  | |α|  | R²   | s(α) |
|----------------|------|---|-----|-----|------|
| **Polish factors** |      |   |     |     |      |
| CAPM           | 8.27 | 0.73 | 45.89 | 0.88 |
| Three-factor   | 1.14 | 0.53 | 59.98 | 0.81 |
| Four-factor    | 1.16 | 0.56 | 60.60 | 0.86 |
| **European factors** |      |   |     |     |      |
| CAPM           | 10.27| 0.63 | 22.44 | 0.89 |
| Three-factor   | 1.65 | 0.62 | 28.36 | 0.89 |
| Four-factor    | 1.70 | 0.64 | 29.23 | 0.89 |
| **Global factor** |      |   |     |     |      |
| CAPM           | 11.24| 0.63 | 24.20 | 0.90 |
| Three-factor   | 1.88 | 0.70 | 29.21 | 0.90 |
| Four-factor    | 2.02 | 0.69 | 30.28 | 0.90 |

Panel B: size and momentum

|                | GRS  | |α|  | R²   | s(α) |
|----------------|------|---|-----|-----|------|
| **Polish factors** |      |   |     |     |      |
| CAPM           | 24.34| 1.25 | 42.20 | 1.39 |
| Three-factor   | 2.78 | 1.13 | 57.09 | 0.97 |
| Four-factor    | 2.08 | 0.87 | 60.48 | 1.05 |
| **European factors** |      |   |     |     |      |
| CAPM           | 21.08| 1.27 | 22.40 | 1.47 |
| Three-factor   | 2.36 | 1.18 | 28.82 | 1.26 |
| Four-factor    | 1.90 | 1.02 | 30.61 | 1.24 |
| **Global factor** |      |   |     |     |      |
| CAPM           | 20.59| 1.26 | 23.64 | 1.48 |
| Three-factor   | 2.60 | 1.19 | 28.88 | 1.32 |
| Four-factor    | 2.37 | 1.09 | 30.97 | 1.31 |

The table shows regression results for the CAPM, three-factor, and four-factor models. The models aim to explain the excess returns of 25 portfolios formed on the book equity to market equity ratio (B/M) and size and 25 portfolios formed on momentum and size. The models’ parameters were estimated based on Polish, European, and global factors. Factors for Europe and the world were obtained from Kenneth French's website. For portfolios created at the end of month t, the lagged momentum return is the stock cumulative return from time $t-12$ to $t-2$. GRS is the Gibbons, Ross, and Shanken [1989] statistic, |α| is the average absolute intercept, R² is the average R², and s(α) is the standard deviation of the intercepts. The critical values for GRS statistics in all models are 1.45 for 90%, 1.61 for 95%, and 1.95 for 99%. Panel A presents the regression results for size and B/M portfolios and panel B refers to size and momentum portfolios. The data period is 11/30/2002–12/31/2013. 

Source: own study.
Finally, as it may be presumed that the 25 portfolios formed from 5 × 5 sorts on size and momentum seem much more difficult with cross-sectional pricing models. First, the performance of the international models is very poor. The models explained only 22–31% of cross-sectional variation and were rejected based on the GRS statistic at the 99% confidence level. Second, the performance of the Polish models based on local factors appears better, but is still far from ideal. The CAPM model explained 42% of cross-sectional variation and the average absolute alpha was 1.25. The GRS statistic of 23.24 definitely rejects the model. The three-factor model reveals some improvement. The R² rose to 57% and the corresponding average absolute intercept fell to 1.13. Further improvement was observed after the inclusion of the momentum factor into the model. The intercept declined to 0.87 and the R² increased to 60%. However, the corresponding GRS statistics for local factors-based three-factor and four-factor models were, respectively, 2.78 and 2.08. In other words, both models are still rejected at the 99% confidence level. Summing up, although the Polish multifactor Fama-French and Carhart models are able to explain the B/M and size sorted portfolios, they fail to fully explain the impact of momentum effect.

Conclusions and Areas for Further Research

The importance of the capital market for Poland’s economy has been growing for the past 20 years. Therefore, there is a need for valid tools to analyze market behavior. The Fama-French and Carhart pricing models are just two examples. The purpose of this paper is to comprehensively examine the applicability of the above-mentioned models on the Polish markets. Achieving this aim relies on several steps, which in themselves may be important for asset management, performance evaluation, and asset pricing.

First, the paper documents the performance of cross-sectional asset pricing factors in Poland. We find particularly strong and statistically significant value and momentum premium. The size premium is also observed, however, it lacks statistical significance. Moreover, interdependencies between factors generally follow patterns observed in developed markets, with the exception of the momentum premium, which seems to be stronger among the large-caps than among the small-caps.

Second, we test the integration of the Polish stock market with international markets. With the exception of momentum, local cross-sectional factors appear to be neither correlated nor explained by European and global factors.

Third, we examine and compare performance of the CAPM, three-factor, and four-factor pricing models in Poland in explaining returns on portfolios double-sorted on B/M and size and on momentum and size. The CAPM model generally fails to explain the value, size, and momentum effect and is rejected. Conversely, the Fama-French and Carhart
models explain well the B/M and size sorted portfolios, with the exception of the most extreme high B/M small-cap portfolio. Finally, all three models fail to fully explain the returns on size and momentum sorted portfolios, although the Carhart model performs best and the CAPM performs worst.

Fourth, we compare the performance of European and global asset pricing factors in explaining cross-sectional variation in returns on the Polish market. In general, all models are rejected. In other words, employing international factors-based models for the Polish market does not seem to be a valid approach.

Further research should concentrate on three main issues. First, it should explore the sources of value, size, and momentum premiums on the Polish market. Second, it should identify the reasons underlying the anomalous behavior of some factors in Poland, which perform differently than in the developed markets (e.g., the reversed large-cap momentum spread is). Finally, whether more sophisticated pricing models, such as the five-factor pricing model [Fama, French, 2013], could also be used for the Polish stocks.

Notes

1 Author’s e-mail address: adam.zaremba@ue.poznan.pl
2 Author’s e-mail address: przemyslaw.konieczka@gmail.com
3 An interesting review may be found in the paper by Karolyi and Stulz [2003].
4 Whenever we used global or the European pricing factors, the analyzed timespan ends in December 2013, as the January data were unavailable when the calculations were performed.
5 In other words, it is a one-year return excluding the last month so as to avoid look-ahead bias.
6 For presentational purposes, we first aggregated cross-sectional arithmetic returns and then computed time-series’ means and standard deviation of quantile portfolios using log-returns.

References


Michou, M., Mouselli, S., Stark, A. (2010), Fundamental analysis and the modelling of normal returns in the UK, working paper, SSRN.


Anchoring and Overconfidence: The Influence of Culture and Cognitive Abilities

Abstract

Anchoring and overconfidence are some of the best-known biases in psychology and behavioral finance literature. While a number of studies have investigated the evidence of these biases and explored the motives and human factors that contribute to one's susceptibility to the effects, little is known about the cultural factors behind these heuristic biases. This paper aims to fill the research gap and shows the differences in proneness to the anchoring effect and overconfidence in two samples of students from Poland and India. The purpose of the study is twofold: to analyze susceptibility to behavioral effects relative to cultural background; and to consider the subjects' cognitive abilities as a potential factor in their exposure to behavioral biases and confirm that subjects with higher cognitive skills, measured by the cognitive reflection test (CRT) display less susceptibility to the above heuristic biases.

Keywords: anchoring, overconfidence, culture, cognitive reflection test
JEL: G02, C91

Introduction

Literature about behavioral finance shows that people often rely on heuristic biases to predict values [Tversky, Kahneman, 1974]. Heuristics are mental shortcuts that allow
Anchoring and Overconfidence: The Influence of Culture and Cognitive Abilities

people to solve problems quickly and make more accurate decisions with less effort [Gigerenzer, 1996]. However, they can also lead to biases in judgments in the condition of uncertainty [Kahneman, Tversky, 1996].

One of the best-known heuristics is anchoring, which implies that individuals judgments often assimilate irrelevant numerical values to which they have been previously exposed. Many studies found undisputed evidence of anchoring in general knowledge [Tversky, Kahneman, 1974; Epley, Gilovich, 2001, 2005], probability estimates [Chapman, Johnson, 1999], legal judgments [Englich, Mussweiler, 2001; Englich et al., 2006], purchasing decisions [Ariely et al., 2003], credit card payments [Stewart, 2009], forecasting [Critcher, Gilovich, 2008], negotiations [Galinsky, Mussweiler, 2001] and self-efficacy [Cervone, Peake, 1986].

The literature review by Furnham and Boo [2011] presents a list of studies that concerns human factors which contribute to anchoring bias susceptibility such as: mood, knowledge/experience, motivation, personality and cognitive ability. However, little is known about the cultural factors behind that heuristic.

Another well-known behavioral bias is overconfidence. Moore and Healy [2008] described it as: (1) overestimation of one’s actual performance, (2) overplacement of one’s performance relative to others, and (3) excessive precision in one’s beliefs. Again, the literature provides evidence of overconfidence in general knowledge [Kahneman, Tversky, 1977; Kahneman, 2011], experimental design [Camerer, Lovallo, 1999; Oskamp, 1965; Klayman et al., 1999], and financial settings [Barber, Odean, 2001; Malmendier, Tate, 2005; Scheinkman, Xiong, 2003; Daniel et al., 2001]. However, susceptibility to overconfidence is not unequivocal and is due to behavioral factors [Kahneman, Tversky, 1977], personality traits [Rzeszutek, 2015], cognitive factors [Hoppe, Kusterer, 2011] or cultural reasons [Antonczyk, Salzmann, 2014; Acker, Duck, 2008].

This study attempts to explain the differences in susceptibility to the anchoring bias and overconfidence among two samples of undergraduate students from Poland and India. The purpose of the study is twofold.

The first is to analyze the roles played by cultural backgrounds. The study of human factors considered in the anchoring literature [Furnham, Boo, 2011] is expanded to include a new human component, namely culture. The traditional assumption of “homo economicus”, a person who is supposed to behave in exactly the same way in Poland and India, is tested. Based on cultural studies (Cheek and Norem, 2016; c), this paper investigates whether the culture has an impact on the anchoring bias and overconfidence. Therefore, the first hypothesis is that subjects coming from different cultures will vary in their susceptibility to the heuristic biases of anchoring and overconfidence.

The second purpose is to consider cognitive abilities as a potential factor in proneness to anchoring bias and overconfidence. Cognitive skills are measured by the cognitive reflection test [Frederick, 2005]. The hypothesis based on Bergman et al. [2010], Oechssler et al. [2009] and Hoppe and Kusterer [2010] states that subjects with higher cognitive
skills display lower susceptibility to the behavioral heuristics and biases such as anchoring and overconfidence.

**Study 1: Anchoring and Overconfidence in a Cultural Dimension**

A substantial body of literature indicates that national culture influences human values [Hofstede et al., 2010; Breuer, Quinten, 2009] and economic outcomes [Guiso et al., 2006; Siegel et al., 2011; Zheng et al., 2012; Frijns et al., 2013; Braginsky, Mityakov, 2015]. Zingales [2015] argues that in the last ten years there has been an explosion of economic research on culture, which originates from the failure of traditional economic models to explain the reality of “homo economicus” embedded in a cultural context. Aggarwal and Goodell [2014] suggest that there is ample opportunity to further investigate the impact of national culture on finance.

Cultural finance literature (as defined by Breuer and Quinten [2009]) deals mainly with differences between Western and Asian cultures. Members of Western societies tend to focus on salient pieces of information and rely more on analytic reasoning. Asians seem to rely more on context and are described as intuitive and holistic thinkers [Nisbett, Masuda, 2003]. Therefore, in the field of behavioral finance, culture started to play an important role in understanding the economic behavior of investors, e.g. the effect of framing, as described in the prospect theory by Kahneman and Tversky [1979], is more noticeable among members of Asian cultures [Wang, Fischbeck, 2004; Levinson, Peng, 2007].

Studies of heuristic biases and cognitive disorders from a cultural perspective show that managers in individualistic countries are characterized by a higher belief in their own abilities and are more susceptible to overconfidence [Antonczyk, Salzmann, 2014; Heaton, 2002; Hackbarth, 2008]. Children who grow up in individualistic countries have the conviction that they are unique, above-average and are born winners. This behavior leads to a strong conviction about their high abilities and infallibility of their estimates [Markus, Kitayama, 1991].

Overconfidence can also affect the investment strategy used by investors. Chui et al. [2010] show that culture can effect stock return patterns and that individualism is positively associated with volume, volatility and profits from momentum strategies. Chui et al. [2010] argue that investors in different cultures interpret information in different ways and are subject to different biases. Investors from less individualistic cultures place lesser weight on information that they come up with by themselves, and more weight on the consensus of their peers and are therefore less prone to overconfidence or self-attribute bias.

Other studies show, however, that Asians are more confident than Britons, although this has more to do with overplacement and the underweighting of the risk of losing [Acker, Duck, 2008; Yates, Lee, 1996; Yates et al., 1996, 1997]. The studies of Jlassi, Naoui and Mansour [2014], which were based on 27 stock exchanges in various cultural regions of the world, confirm that Asian markets (including Hong Kong, India and the
Philippines) experience the highest level of price changes and that Asian investors are subject to overconfidence.

One of the most popular and most common heuristic biases – anchoring – was also analyzed. Studies by Cheek and Norem [2016] sought to verify the hypothesis that people who think holistically should be more affected by anchoring than those who think analytically. Holistic thinkers are considered inter-dependent because they focus largely on the context instead of specific elements, and therefore should be subject to anchoring to a greater extent. Although individuals characterized by analytic thinking consider themselves independent, they should be more suspicious with regards to the anchor given [Choi et al., 2007; Nisbett et al., 2001]. Empirical studies by Cheek and Norem [2016] have not, however, confirmed this hypothesis and show that people characterized by a holistic way of thinking are less susceptible to heuristic anchoring than those characterized by analytical thinking.

Study 2: Anchoring and Overconfidence – The Influence of Cognitive Abilities

The experimental study outlined in this section examines the extent cognitive skills play a role in decision-making behavior involving heuristics, especially anchoring and overconfidence.

Due to bounded rationality, people are unable to make fully logical decisions. Kahneman and Tversky [1984] state that cognitive information is processed in our brain by system 1 or system 2. System 2 requires us to conduct effortful, demanding and reflective mental activities while system 1 operates quickly, automatically, intuitively and without requiring a lot of time or effort.

The Cognitive Reflection Test (CRT) proposed by Frederick [2005] is a 3-item task that measures the extent to which individuals form their judgments intuitively, as opposed to through reflection. The CRT is designed to measure the tendency to override a prepotent response alternative that is incorrect and to engage in further reflection that leads to a correct answer [Toplak et al., 2011]. Studies showed that CRT predicts the susceptibility to decision-making biases and heuristics better than intelligence tests [Toplak et al., 2011, 2014] and is not just a mathematical test but measures something above and beyond general skills, namely cognitive reflection [Campitelli, Gerrans, 2013]. Pennycook et al. [2016] provide the evidence that CRT is more a measure of reflective than intuitive thinking.

The CRT has been used in several studies to measure the effect of cognitive abilities on the susceptibility to behavioral biases [Oechssler et al., 2009; Duttle, Inukai, 2015; Toplak et al., 2011]. In almost all cases, the higher CRT scores [Frederick 2005] were correlated to lower incidences of analyzed heuristics or behavioral biases.

Two recent studies explored the case of anchoring and cognitive ability. Bergman et al. [2010] showed that the anchoring effect decreased, but did not vanish, with a higher cognitive ability measured by a cognitive ability test (CAT) and the CRT. However, they suggested that CAT is a better predictor of anchoring than the CRT instrument. Oechssler
et al. [2009] also tested some behavioral biases in the context of cognitive abilities with the CRT, but found that test scores had no influence on the degree of anchoring.

Hoppe and Kusterer [2011] explored the issue of overconfidence and cognitive ability using a CRT. They showed that more intuitive decision makers are relatively less successful in assessing the right number of correct answers to questions related to general knowledge, although there was no clear tendency indicating that they are more overconfident than analytical decision makers.

**Experimental Design**

Students from two countries (Poland and India) were studied between 2014–2015. The sample groups were selected to represent two different cultural regions: Eastern Europe and Southern Asia.

The Polish respondents in the study included students from the Warsaw School of Economics. The group comprised 77 people. In the studied sample group, there were 40 women and 37 men. They were mainly Finance and Accounting majors. The group of students from India from the Management Development Institute in Gurgaon comprised 66 MBA program students, comprising 31 women and 33 men.

The study had a quasi-experimental (comparison) character. With regards to classification variables like gender or culture, one has to make use of the existing levels of the respective variable. In this scheme, one is able to study the difference between groups, although it is not possible to randomly select the study groups, because the author measured and compared naturally distinct groups. Consequently, the non-probability technique of selecting samples was random sampling, based on convenience sampling. This selection method was also chosen due to the considerable difficulties and high costs associated with using other techniques of sample selection when studying foreign respondents.

The study relied on a questionnaire composed of questions concerning the extent of anchoring effect and overconfidence (overprecision and overplacement). The drafting of questions in the questionnaire was based on the work by Jacowitz and Kahneman [1995], Soll, Klayman [2004], Larrick et al. [2007]. Questions were intentionally simple and referred to universal objects (height of the tallest tree in the world) and neutral opinions (American debt, personal evaluation of one’s ability) as to avoid a lack of understanding between two groups of students (see Appendix 1, question 1, 2, 3). In order to measure cognitive ability, the three-item cognitive reflection test (CRT) introduced by Frederick [2005] was used (see Appendix 1, question 4, 5, 6).

**Differences Between Polish and Indian Sample Based on the Model of Hofstede**

According to Hofstede et al. [2010], 4 fundamental cultural dimensions can be distinguished: the individualism dimension (IDV), power distance (PDI), uncertainty
avoidance (UAI) and masculinity (MAS). Dimensions for Poland and India are presented in Table 1.

<table>
<thead>
<tr>
<th>TABLE 1. Dimensions of Hofstede’s cultural model for: Poland and India</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>Poland</td>
</tr>
<tr>
<td>India</td>
</tr>
</tbody>
</table>

Source: own study based on: Hofstede et al. [2010].

The biggest difference between Poland and India concerns the uncertainty avoidance dimension. Poland is a country with one of the higher UAI indicators, while India is a collectivist country with an average UAI. Citizens of countries with a high UAI index are, on average, more emotional, express more social concern and, in general, have a higher aversion to risk. On the other hand, countries with a low UAI index accept and cope better with unstructured situations, changes in the social environment, and prefer less regulation. Citizens of these countries seem to be more pragmatic and have a better tolerance to change. They are also characterized by a lower aversion to risk.

Cultural Dimensions and Anchoring

Respondents had to state what they thought the height of the tallest tree in the world was and were split into two groups. The first were asked the question with an anchor of 55 meters, and the second with an anchor of 365 meters (see Appendix 1, question 1). The drafting of the question was based on the work by Jacowitz and Kahneman [1995, p. 1163] “height of the tallest redwood”.

The aim of the study was to examine to what degree the students’ responses would be close to the anchoring values stated. The anchoring indicator was calculated using the following formula (Jacowitz and Kahneman [1995], Kahneman [2011]):

\[
\text{Anchoring indicator} = \frac{\text{difference between average estimates}}{\text{difference between the anchors}}.
\]

According to Kahneman [2011], the typical value of this indicator fluctuates around 55%. A high value of this indicator means that the test subjects were affected by the studied heuristic bias to a large extent. If people slavishly cling to the anchoring values, the anchoring indicator is 100 percent. If people were able to completely ignore the anchoring value, it would amount to 0 percent.
**Cultural Dimensions and Overconfidence – Overprecision**

The study was subject to miscalibration in the form of overconfidence about the precision of information held – also known as overprecision. Respondents had to estimate the level of American public debt as per 04.07.2014. They were asked to state a range about which they were 98% certain (see Appendix 1, question 2). The question was based on the work by Soll, Klayman [2004], where participants were asked to provide interval estimates.

The task is assessed positively if the respondent specifies such a wide range that the right answer to the question is between the stated brackets. The variable of overconfidence was the precision of the forecast. If the forecast value was within the range, 1 point was awarded, otherwise no points were granted. Subsequently percentages were calculated for the correct and incorrect estimates for each attempt and country. The variable of overconfidence at aggregate level equaled the percentage of incorrect estimates (incorrect ranges).

**Cultural Dimensions and Overconfidence – Overplacement**

The study considered the inclination to assume one is a better person in a particular area than the average unit, overplacement, by asking typical questions about the chances of the person who was the subject of the study completing their course with an above-average grade (see Appendix 1, question 3). The drafting of the question was based on the work by Larrick et al. [2007]. In this case, overplacement meant that the majority of subjects replied positively to the above question, while de facto more than half respondents cannot be above average.

**Results**

**Anchoring – Influence of Culture**

In order to check whether the students were affected by the anchoring, and (if so) whether the degree to which they were affected differed by country, an inter-group comparison was performed using the Mann-Whitney U test. In the analyses performed, the group variable was the amount of the anchoring value (365 m versus 55 m), and the dependent variable was the height of the tallest tree in the world estimated by students. The comparison was performed separately in groups of students from Poland and India. The results are summarized in Table 2.

This indicates that estimations by both groups about tree height differed to a statistically-significant level, depending on the anchoring value. Where the anchor was 365 m, the average estimated height of the tree in all student groups was greater than in groups where the anchor was 55 m. The values of the anchoring indicator show that students from Poland were less affected by heuristic bias (anchoring value = 35.5%) than students from India (anchoring value = 72.8%).
TABLE 2. The effect of anchoring values on the estimation of the height of the tallest tree in the world, based on nationality

<table>
<thead>
<tr>
<th>Country</th>
<th>Anchor: 365 m</th>
<th>Anchor: 55 m</th>
<th>Mann-Whitney test</th>
<th>Anchoring indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>M</td>
<td>SD</td>
<td>n</td>
</tr>
<tr>
<td>Poland</td>
<td>38</td>
<td>195.55</td>
<td>157.75</td>
<td>33</td>
</tr>
<tr>
<td>India</td>
<td>27</td>
<td>340.07</td>
<td>271.85</td>
<td>27</td>
</tr>
</tbody>
</table>

Source: own study.

Overprecision – Influence of Culture

To check whether the propensity to overprecision is due to cultural conditioning, a chi-squared test for independence was performed. In the analysis, interdependence between the examined countries and percentage of students affected by heuristic bias were studied. The results of the test show a statistical tendency $\chi^2(df = 1) = 3.10; p = 0.078$. It can be cautiously assumed that the proportion of people affected by overprecision is not the same in the compared groups. The percentage of Indian students affected by overprecision was similar at 85.7% ($n = 42$); for students from Poland, the percentage was lower, 71.7% ($n = 43$). Figure 1 shows the difference between students with regards to overprecision.

FIGURE 1. Overprecision of students from Poland and India

Source: own study.

The study also examined whether the proportion of people undervaluing, giving the correct answer and those overvaluing differed depending on the country of study (Figure 2). The results of the chi-squared test are not statistically significant, $\chi^2(df = 2) = 3.22; p = 0.200$. However, it can be observed that the vast majority of people from India gave
a lower range of values, i.e. they underestimated the value. This could be related to the phenomenon of underestimating the risk of loss [Acker, Duck, 2008; Yates, Lee, 1996; Yates et al., 1996, 1997].

**FIGURE 2.** Overprecision of students from Poland and India (under- and overestimation)

![Graph showing overprecision of students from Poland and India](image)

*Source: own study.*

**Overplacement – Influence of Culture**

To check whether the tendency to have excessive confidence in oneself (overplacement) is based on cultural factors, the author performed an analysis using a chi-squared test for independence. In this analysis, interdependence between the examined country and the percentage of students affected by overplacement was studied. The results of the test were not statistically significant, $\chi^2(df = 1) = 0.01; p = 0.938$. There are no grounds to state that the level of overplacement differs depending on the nationality of the students. In the groups from Poland and India, the percentage of students expecting to complete their studies with an above-average grade was similar, at 89.5% ($n = 68$) and 89.1% ($n = 57$) respectively.

**Anchoring and the CRT (Without a Division by Country)**

The analysis considered the influence of anchoring on the estimation of the size of an object, based on the results of the cognitive reflection test. For this purpose, the Mann-Whitney $U$ test was used to compare the estimated height of a tree with anchors of 365 m and 55 m, among all student groups (from Poland and India) with both positive and negative results from the CRT. The answers of all respondents were mixed in order to exclude cultural factors from the analysis and to check exclusively for the influence of cognitive indicators on the degree of susceptibility to the particular heuristic bias. The results of the analysis are summarized in Table 3.
TABLE 3. Influence of anchoring on the estimation of the height of a tree, based on the results of the CRT

<table>
<thead>
<tr>
<th>CRT score</th>
<th>Anchor: 365 m</th>
<th>Anchor: 55 m</th>
<th>Mann-Whitney test</th>
<th>Anchoring indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>M</td>
<td>SD</td>
<td>n</td>
</tr>
<tr>
<td>Positive</td>
<td>27</td>
<td>236.74</td>
<td>201.60</td>
<td>25</td>
</tr>
<tr>
<td>Negative</td>
<td>35</td>
<td>284.89</td>
<td>256.68</td>
<td>35</td>
</tr>
</tbody>
</table>

Source: own study.

The results of the Mann-Whitney U test show that estimates based on data from students from Poland and India differ to a statistically significant extent, depending on the value of the anchor. In cases where the anchor was 365 m, the average height estimate for the tree was higher than in cases where the anchor was 55 m. The values of the anchoring indicators show that heuristics bias has less influence on students with a positive score in the CRT. The value of the anchoring indicator in this group was 46.0%. For students with a negative score in the CRT, the value of the indicator was higher at 58.3%. This means that if cultural factors are excluded from the analysis, the results of the CRT, cognitive abilities, have an influence on the degree of proneness to the respective heuristic bias. This confirms the hypothesis of Toplak, West and Stanovich [2011], which states that the CRT is, on the one hand, a measure of rational thinking, and on the other hand the degree of susceptibility to cognitive distortions.

Anchoring and the CRT (with a Division by Country)

Next, I analyzed the influence of anchoring on estimating the size of an object, based on the results of the cognitive reflection test divided by country of origin. Here, the Mann-Whitney U test was used to compare the estimated tree height with anchors of 365 m and 55 m, amongst student groups with both positive and negative results from the CRT. The analysis was performed separately for students from Poland and India. The results of the analysis of the student from Poland are shown in Table 4, and for the student from India in Table 5.

TABLE 4. Influence of anchoring on the estimation of the height of a sequoia, based on the results of the CRT – students from Poland

<table>
<thead>
<tr>
<th>CRT score</th>
<th>Anchor: 365 m</th>
<th>Anchor: 55 m</th>
<th>Mann-Whitney test</th>
<th>Anchoring indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>M</td>
<td>SD</td>
<td>N</td>
</tr>
<tr>
<td>Positive</td>
<td>16</td>
<td>163.44</td>
<td>97.51</td>
<td>17</td>
</tr>
<tr>
<td>Negative</td>
<td>21</td>
<td>223.62</td>
<td>192.27</td>
<td>16</td>
</tr>
</tbody>
</table>

Source: own study.
For students from Poland the value of the anchor had a significant influence on the estimated height of the tree. Both in the group with positive and negative results in the CRT, the average estimated height of the tree in cases where the anchor was 365 m was greater than in cases where the anchor was 55 m. At the same time, the anchoring indicator in the group of people with a negative score in the CRT was higher than in the group of people with a positive score in the CRT.

**TABLE 5. Influence of anchoring on the estimation of the height of a sequoia, based on the results of the CRT – India**

<table>
<thead>
<tr>
<th>CRT score</th>
<th>Anchor: 365 m</th>
<th>Anchor: 55 m</th>
<th>Mann-Whitney test</th>
<th>Anchoring indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>M</td>
<td>SD</td>
<td>N</td>
</tr>
<tr>
<td>Positive</td>
<td>19</td>
<td>298.47</td>
<td>245.04</td>
<td>18</td>
</tr>
<tr>
<td>Negative</td>
<td>6</td>
<td>499.33</td>
<td>351.81</td>
<td>9</td>
</tr>
</tbody>
</table>

*Source: own study.*

For students from India the average estimated tree height when the anchor was 365 m was higher, both for people with a positive and negative score in the CRT. However, due to the low number of people with a negative score in the CRT, the difference between estimations where the anchor was 365 m and 55 m is only at the level of a statistical tendency \( p = 0.098 \). The values of the anchoring indicators show that people with a negative score in the CRT are more susceptible to anchoring than people with a positive score.

The analysis shows that both cognitive and cultural indicators have an effect on the degree of exposure to the heuristic bias.

**Overconfidence and the CRT (Without a Division by Country)**

Analysis of the link between the score in the CRT and overprecision follows. Using the chi-squared test, I checked whether the proportion of people susceptible to this heuristic bias was the same in the group of students with positive and negative scores in the CRT. The results of the chi-squared test were also insignificant statistically \( \chi^2(df = 1) = 0.03; p = 0.859 \). There are therefore no grounds to conclude that the variables are linked. The proportion of people susceptible to overprecision was similar, regardless of the results of the cognitive reflection test. Figure 3 shows the studied dependency.

In the next stage of the analysis, the link between the results of the CRT and overplacement was examined. The chi-squared test was used to check whether the proportion of people susceptible to this heuristic bias was the same in the group of students with positive and negative CRT results. The outcomes of the chi-squared test are insignificant statistically \( \chi^2(df = 1) = 0.15, p = 0.700 \). There are therefore no grounds to conclude that the variables
are linked. The proportion of people prone to overplacement was similar, regardless of the results of the cognitive reflection test. Figure 4 shows the studied dependency.

FIGURE 3. Overprecision based on the results of the CRT

![Bar chart showing overprecision based on CRT results](chart1)

Source: own study.

FIGURE 4. Overplacement based on the results of the CRT

![Bar chart showing overplacement based on CRT results](chart2)

Source: own study.

Overconfidence and the CRT (with a Division by Country)

In the next stage of the analysis, the link between the results of the CRT and overprecision was studied. The chi-squared test was used to check whether the proportion of people susceptible to this heuristic bias was the same in the group of students with positive and
negative results in the CRT. The analysis was performed separately for students from Poland and from India. The results of the chi-squared test are shown in Table 6. The proportion of people susceptible to overprecision was similar for students from Poland and India, regardless of the results of the cognitive reflection test.

### TABLE 6. Overprecision based on the results of the CRT, by country

<table>
<thead>
<tr>
<th>Country</th>
<th>CRT score</th>
<th>Chi-squared test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>positive</td>
<td>negative</td>
</tr>
<tr>
<td>Poland</td>
<td>69.7% ($n = 23$)</td>
<td>74.1% ($n = 20$)</td>
</tr>
<tr>
<td>India</td>
<td>84.2% ($n = 32$)</td>
<td>91.1% ($n = 10$)</td>
</tr>
</tbody>
</table>

Source: own study.

The link between the results of the CRT and overplacement was also examined. The chi-squared test was used to check whether the proportion of people affected by this heuristic bias was the same in groups of students with positive and negative CRT scores. The analysis was performed separately on students from Poland and India. The results of the chi-squared test are shown in Table 7. The results obtained indicate that there is no basis to believe that the variables are linked. The proportion of people vulnerable to overplacement was similar for students from Poland and India, regardless of the results of the cognitive reflection test.

### TABLE 7. Overplacement based on the results of the CRT, by country

<table>
<thead>
<tr>
<th>Country</th>
<th>CRT score</th>
<th>Chi-squared test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>positive</td>
<td>negative</td>
</tr>
<tr>
<td>Poland</td>
<td>94.6% ($n = 35$)</td>
<td>84.2% ($n = 32$)</td>
</tr>
<tr>
<td>India</td>
<td>86.0% ($n = 37$)</td>
<td>95.0% ($n = 19$)</td>
</tr>
</tbody>
</table>

The percentages represent the amount of people susceptible to excessive confidence. Source: own study.

### Discussion of Findings and Conclusion

The statistical analysis performed in study 1 revealed that respondents from the two studied groups were affected by anchoring, and the degree to which they were influenced by the respective heuristic bias differed depended on their country of origin. The values of the anchoring indicator show that students from Poland (35.5%) are less susceptible
to heuristic bias than students from India; the value of the anchoring indicator was twice as high at 72.8%. The higher anchoring indicator for students from India, compared to students from Poland, confirms the hypotheses of Choi, Koo & Choi [2007] as well as Nisbett, Peng, Choi & Norenzayan [2001], which states that people who think in a holistic manner are more susceptible to the phenomenon of anchoring and are therefore more prone to context manipulation relating to aspects of the situation.

Our statistical analysis also showed key statistical differences in the susceptibility to overconfidence, depending on the country of respondent’s origin. In the case of miscalibration due to overprecision, the percentage of people with overconfidence among students from India was 85.7%. For students from Poland the percentage was 71.7%. The Polish students, characterized by a higher indicator of individualism than students from India, were less exposed to overconfidence. This did not confirm the results of the tests by Antonczyk, Salzmann [2014], Heaton [2002] Hackbarth [2008], Markus and Kitayama [1991] or Chui, Titman, Wei [2010], who showed evidence of a positive dependence between the indicator of individualism and overconfidence. An in-depth analysis of the structure of responses showed, however, that the higher confidence from students in India was primarily due to an underestimation of the estimated results, which would confirm the hypothesis of Acker and Duck [2008], who consider the underweighting of downside risk by Asian cultures. This is confirmed by an analysis of the Hofstede indicators [Hofstede et al., 2010], which shows that the UAI scores for India are amongst the medium and low scores internationally, indicating that these societies are less averse to risk-taking than, for example, Poland (which has one of the highest scores for UAI worldwide).

When examining the tendency to believe in being above-average in a particular area (overplacement), it was shown that the level of overconfidence did not differ to a statistically significant degree by student nationality. In both groups the percentage of students expecting to complete their studies with an above-average grade was similar at 89.5% and 89.1% respectively.

The statistical analysis performed in study 2 verified the hypothesis that subjects with higher cognitive skills display weaker proneness to behavioral heuristics. The hypothesis was confirmed with regards to anchoring. Both at the aggregate and individual level the author shows that people with a negative score in the CRT are more susceptible to anchoring than people with a positive score. The study results confirm the findings of Toplak et al. [2011, 2014] and Bergman et al. [2010] stating that the results of the CRT are a good forecast tool concerning susceptibility to the heuristic biases.

Clearly, the sample size imposed some limitations on the study, since it was not possible to perform a regression analysis to check whether variables (cultural or cognitive) have a greater effect on the susceptibility to the respective heuristic bias. A second limitation was usage of the Cognitive Reflection Test with open-ended questions as in Kahneman [2011, p. 65]. Such a modification of the CRT could, on the one hand, slightly enhance the
overall CRT results of the students, but on the other, hand eliminate potential mathematical errors mentioned by Pennycook et al. [2016].

In the future analyses, the author hopes to increase the sample size, expand it to cover other countries from different cultural regions, and explore the role of gender or religion [Czerwonka, 2014] and expand the analysis by CRT 7 [Toplak et al., 2014].

Notes

1 Author’s e-mail address: monika.czerwonka@sgh.waw.pl

References

Anchoring and Overconfidence: The Influence of Culture and Cognitive Abilities


Appendix 1

1. Is the height of the tallest tree in the world more than 55 meters /365 meters (second group)
   ☐ Yes  ☐ No
   In your opinion, what is the height of the tallest tree in the world? …

2. Estimate the level of American public debt in USD as per 20.05.2014 (state the range within which the right answer should be in order to have 98% certainty)
   …………………….$ – …………………….$

3. Do you think you can finish your studies with the better score than the average?
   ☐ Yes  ☐ No

4. A bat and a ball cost $ 1.10 in total. The bat costs $ 1.00 more than the ball. How much does the ball cost?
   ☐ a) $ 1,  ☐ b) $ 0.05,  ☐ c) $ 0.10,  ☐ d) $ 0.01

5. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?
   ☐ a) 500 minutes  ☐ b) 100 minutes  ☐ c) 50 minutes  ☐ d) 5 minutes

6. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?
   ☐ a) 48 days  ☐ b) 24 days  ☐ c) 47 days  ☐ d) 42
Culture-Based Rejection of Taboo-Infringing Imports

Abstract

This empirical study analyzes the cultural basis of the United States market response to imported Spanish products that seem to violate strongly-held cultural taboos. Survey responses were obtained from students in two contrasting majors, Art and Business, in two distinct cities and universities, i.e. Little Rock at the University of Arkansas, and Dominguez Hills at California State University. The study focused on a baby doll marketed to piggy-back on the new movement towards breastfeeding babies. Although accepted in its original European market, the United States media reports strong moral objections to this product among U.S. citizens. The toy was overwhelmingly rejected in some, but not all, population sub-groups. This study attempts to discern the cultural basis for product rejection by comparing responses between regions, college majors, genders and gender/major combinations. Differences in acceptance between groups are correlated with specific cultural constructs.
Keywords: culture-based, import rejection, taboo, taboo-infringing, international marketing, Spain, USA, culture dimensions, gender, International Journal of Globalization and Small Business

JEL: F18, F19, M30, M31, M39, Y9, Z10, Z13, Z19

Introduction

This research analyzes U.S. market acceptance and non-acceptance of an apparently culture-laden import from Spain. The product is a life-sized baby doll called Bebe Gloton in Spain and Breast Milk Baby in the U.S., which is produced and distributed by Berjuan, S.L. of Alicante, Spain. Over two million Bebe Gloton dolls were sold in Europe [Peaceful Parenting, 2009] and the product was introduced to the U.S. market in 2011 as Breast Milk Baby in male, female, and three different ethnic versions [Chicago Tribune, 2011]. The doll’s introduction created significant a media controversy in the U.S, characterized by reports of extreme aversion to the product. TV-news correspondent Bill O’Reilly even denounced the doll on Fox News [CBS News, 2012]. This American response, in contrast to overall acceptance in Europe, indicates the existence of a cultural taboo in the U.S. [Tews, Halliburton, 2014].

One could assume that common products widely used worldwide, such as traditional dolls, would not be especially prone to violating cultural taboos. Most baby-doll-related negative reactions in the U.S. have been focused on ethnic issues. This may be the first baby-doll product to elicit a reaction related to sexuality. Importing culture-laden products into culturally different markets may, however, easily violate cultural taboos. Taboos are not generally talked or written about, and thus may be invisible to the importer until after market introduction. This research attempts to quantify the cultural-basis, if any, underlying market reaction to this product in (i) two U.S. regions, (ii) between genders, (iii) between Art and Business majors, and (iv) between the four gender/major combinations. These groups have previously been found to exhibit significant differences [Brice, 2016]. Responses will be analyzed to find correlations between product acceptance and scores for seven specific culture constructs. Thelen, et. al. [2003] found that although antecedents of consumer ethnocentrism may be significant for a nation as a whole, they are most meaningful in some, but not all sub-cultures. This implies that there may be critical differences in product acceptance or non-acceptance between different geographic regions or sub-groups and also different specific cultural traits. Such information should help marketers who selling culture-laden products to better chart these differences.
Cultural Taboos and Product Acceptance

The global dissemination of the U.S. products and ideas has, in some countries, replaced traditional patterns with the American lifestyle patterns. The process underlying the spread of non-American culture into the U.S. should be similar [Craig et. al., 2009]. Efforts by foreign (non-U.S.) firms should gradually result in the introduction of foreign products and ideas that may eventually become an integral part of the US culture.

Hofstede [2001] states, “Culture is the collective programming of the mind that differentiates groups of people”. Americanization is interpreted in some countries as an embodiment of globalization and is criticized in others as an attack on traditional culture [Craig et. al., 2009]. Products impact culture because they contain cultural meaning [McCracken, 1986] and can be viewed as culture-laden. The meaning of products is linked to ingrained values and beliefs. Culture is contained and interpreted in the values and norms governing objects used in everyday life and, in marketing, attention is focused on culture’s impact on buying behavior. Product ownership may define lifestyle membership, as well as culture’s values, attitudes and beliefs. Cultural norms shape societal attitudes, lifestyles and products which are embedded in that culture, and also inhibit the acceptance of opposing attitudes, lifestyles and products foreign to that culture [Craig et. al., 2009].

Taboo advertising involves products and ideas that have a culturally forbidden status, and their advertising may cause offence by violating cultural norms [Sabri, 2012a, 2012b]. Some products are more objectionable than others but offensive ideas in advertising often shock more than the products themselves [Beard, 2008]. Taboo product research has shown that taboo-violating advertising can actually be completely unacceptable [Fahy et al., 1995; Waller, 1999]. Because of the power contained in taboo-violating advertisements, marketers increasingly use them as a cost effective way to gain recognition [Pope et. al., 2004].

Researchers have found the optimum amount of marketing taboo-violation that attracts maximum attention; and that humor can decrease taboo and increase brand remembrance [Cline, Kellaris, 2007; Madden, Weinberger, 1982]. Overuse of a taboo tactic becomes counterproductive [Sabri, 2012b]. Because taboos are cultural, different cultures will respond differently to the same product addressing the same taboo [Cheng, 2009]. Parry et al. [2013] also found that reactions differed by religion and gender.

Measuring Market Culture

International marketing literature demonstrates that behavior based on culture is strongly related to product acceptance in different markets. Cultural constructs have been developed to measure different aspects of culture. Most of these constructs examine
national-level culture, but some constructs have been developed in order to be used at the organizational or professional-field level.

Five constructs measuring beliefs – called social axioms – have been created. They can be used to distinguish segments of society at the national and sub-national level [Leung et al., 2002]. These axioms are Social Cynicism, Social Flexibility, Reward for Application, Spirituality, and Fate Control. Social Cynicism involves a negative attitude towards society and manipulation. Social Flexibility includes the belief that behavior may sometimes be contradictory, depending on the situation. Reward for Application is the belief that constant effort will have positive outcomes. Spirituality entails the supernatural. Fate Control involves an attitude towards predestination and peoples’ impact on future events.

Geert Hofstede [1980, 1991, 2001] created a number of cultural dimensions based on values, two of which have been shown to work for sub-national groups. Power Distance concerns the nature of unequal power in society and its acceptance or non-acceptance. Social classes, and their related occupations, have separate cultures and findings significantly differ according to occupation. The highest scores were related to the lowest status and countries with the lowest Power Distance scores had the largest spread of occupational scores [Hofstede, 1991]. The Masculinity vs. Femininity dimension relates to value differences between genders. In high Masculinity cultures, high scores relate to high differentiation in gender roles as well as strong concern among males for materialism. In high Masculinity cultures women score significantly lower than men. In low Masculinity cultures (high Femininity) scores of both genders are similar or identical. Both genders rate quality of life issues as more important than pure materialism [Hofstede, 2001].

The research developed by Hofstede and Leung, Bond et al. is seminal to the field of culture in the context of international management. The combined cultural constructs they have developed are established and widely disseminated, and have been repeatedly found to be statistically valid. They cover both values and beliefs that do not have any strong overlap in definition. For these reasons, and especially because of the wide range of unrelated cultural constructs they cover, these seven constructs have been chosen for this study.

Hypotheses

Previous research showed a significant correlation between specific cultural constructs and business performance outcomes [Brice, 2013]. It has been found that reflection of the Bebe Gloton doll was significantly different between Art and Business majors, gender and gender/major combinations [Brice et al., 2016]. It is expected then that differences among gender and major reflect specific differences in cultural values and beliefs and would show significant correlation with specific cultural constructs. The following relationships are
expected due to the previous research implicating progressive cultural values with greater rejection, and conservative cultural values with greater acceptance [Brice et al., 2016].

H1: The choice to allow a child to Play with the Bebe Gloton doll would significantly correlate with one or more of the seven cultural constructs.
H1a: PDI would show a significant positive correlation with Play.
H1b: MAS would show a significant positive correlation with Play.
H1c: Spirituality would show a significant positive correlation with Play.
H1d: Social Cynicism would show a significant positive correlation with Play.
H1e: Reward for Application would show a significant positive correlation with Play.
H1f: Fate Control would show a significant positive correlation with Play.
H1g: Social Flexibility would show a significant negative correlation with Play.

H2: Major would significantly correlate with one or more of the seven cultural constructs.
H2a: PDI would show a significant correlation with Business.
H2b: MAS would show a significant correlation with Business.
H2c: Spirituality would show a significant correlation with Business.
H2d: Social Cynicism would show a significant correlation with Art. H2e: Reward for Application would show a significant correlation with Art. H2f: Fate Control would show a significant correlation with Business.
H2g: Social Flexibility would show a significant correlation with Art. H3: Gender would significantly correlate with one or more of the seven cultural constructs.

H3a: PDI would show a significant correlation with Male.
H3b: MAS would show a significant correlation with Male.
H3c: Spirituality would show a significant correlation with Male.
H3d: Social Cynicism would show a significant correlation with Male.
H3e: Reward for Application would show a significant correlation with Male.
H3f: Fate Control would show a significant correlation with Male.
H3g: Social Flexibility would show a significant correlation with Female.

H4: Major-Gender groups, as a single variable, would significantly correlate with one or more of the seven cultural constructs.

H5: One or more specific Major-Gender pairings would significantly differ with each other on one or more of the seven cultural constructs.

H6: One or more specific Major-Gender pairings would significantly differ with each other with respect to the variable Play.
**Methods**

A survey instrument was administered to 242 respondents on two college campuses; the University of Arkansas at Little Rock (UALR) and California State University Dominguez Hills (CSUDH). Data were collected in-person and on-campus at the Business and Art departments. 144 Business students and 89 Art students provided complete responses: 45 Business students at UALR and 99 at CSUDH; 41 Art students at UALR and 48 at CSUDH. The study involved 91 Male and 142 Female students: 30 Male students from UALR and 61 from CSUDH; 56 Female students from UALR and 86 from CSUDH (see Table 1).

<table>
<thead>
<tr>
<th>TABLE 1. Respondent Demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>UALR</td>
</tr>
<tr>
<td>CSUDH</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

*Source: own study.*

The survey instrument featured a photo of the Bebe Gloton doll product in its original packaging, with a description of the product’s attributes and intended use. Part 1 of the survey asks demographic questions and studies respondents’ reactions to the product (see survey in Appendix).

Questions from part 1 examined whether there would be anything “immoral” about the product if it were intended for a 5-year-old girl; if the respondent would allow their presumed 5-year-old girl to “play” with the product; and finally if they would “buy” the doll. The variable Play was chosen to be the measured dependent variable because it was the response with the least restrictive interpretation.

Part 2 questions represented five Leung and Bond cultural constructs (Social Cynicism, Social Flexibility, Reward for Application, Spirituality, and Fate Control), with two Hofstede constructs (Power Distance and Masculinity-Femininity).

**Results**

H1-H1g: The variable Play significantly correlated with PDI (Pearson Chi-Square of 0.008). Play did not correlate with any other construct. Because one of the seven cultural constructs significantly correlated with Play, H1 is supported. H1a is supported in that the relationship between Play and PDI is positive instead of negative. Thus, the choice of
Play is more likely when PDI is high. H1b through H1g are not supported as there were no other significant correlations.

H2-H2g: College majors significantly correlated with MAS (Pearson Chi-Square of −0.038), Spirituality (Pearson Chi-Square of −0.014), and Fate Control (Pearson Chi-Square of −0.004). All three were positively correlated with Business as expected. The major Business is associated with higher Masculinity, higher Spirituality, and higher Fate Control. Thus H1, H2b, H2c, and H2f were supported while H2a, H2d, H2e, and H2g were not supported with significant findings.

H3: Gender significantly correlated with Social Cynicism (Pearson Chi-Square of −0.008), thus supporting H3 and furthermore correlated with Male, supporting H3d. However, H3a, H3b, H3c, H33, H3f, and H3g were neither supported nor reversed with significant findings.

H4: The four Major-Gender groups, when treated as a single variable, significantly correlated with PDI (Kruskal Wallis of 0.032). Because no other cultural construct showed significance Power Distance may be considered the primary cultural construct difference among these four groups, and H4 is supported (see Table 2).

### TABLE 2. Major-Gender Grouping vs. Cultural Construct (Kruskal Wallis)

<table>
<thead>
<tr>
<th>Major-Gender Groupings*</th>
<th>Cultural Construct</th>
<th>Kruskal Wallis Significance</th>
<th>Average Construct Scores (0–5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BM, BF, AM, AF</td>
<td>Power Distance</td>
<td>0.032</td>
<td>Bus M 3.53, Art M 3.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bus F 3.30, Art F 3.25</td>
</tr>
<tr>
<td>AM &amp; BF</td>
<td>Spirituality</td>
<td>0.017</td>
<td>Bus F 3.18, Art M 2.98</td>
</tr>
<tr>
<td>AF &amp; BM</td>
<td>Power Distance</td>
<td>0.014</td>
<td>Bus M 3.53, Art F 3.25</td>
</tr>
<tr>
<td>AM &amp; BM</td>
<td>Spirituality</td>
<td>0.032</td>
<td>Bus M 3.17, Art M 2.98</td>
</tr>
<tr>
<td>AM &amp; BM</td>
<td>Reward for Application</td>
<td>0.049</td>
<td>Bus M 3.99, Art M 3.82</td>
</tr>
<tr>
<td>AF &amp; BF</td>
<td>Social Cynicism</td>
<td>0.014</td>
<td>Art F 3.20, Bus F 2.84</td>
</tr>
<tr>
<td>AM &amp; AF</td>
<td>Masculinity</td>
<td>0.040</td>
<td>Art M 3.89, Art F 3.61</td>
</tr>
<tr>
<td>AM &amp; AF</td>
<td>Spirituality</td>
<td>0.044</td>
<td>Art F 3.21, Art M 2.98</td>
</tr>
<tr>
<td>BM &amp; BF</td>
<td>Power Distance</td>
<td>0.013</td>
<td>Bus M 3.53, Bus F 3.30</td>
</tr>
<tr>
<td>BM &amp; BF</td>
<td>Fate Control</td>
<td>0.028</td>
<td>Bus M 2.67, Bus F 2.36</td>
</tr>
</tbody>
</table>

* A = Art, B = Business, M = Male, F = Female

Source: own study.

H5: Nine specific Major-Gender pairings significantly differed from each other with respect to various cultural constructs, including PDI, Masculinity, Spirituality, Reward
for Application, Social Cynicism, and Fate Control. Social Flexibility did not significantly correlate with any pairing. H5 is thus supported (see Table 2).

H6: The Art/Male and Business/Female pairing significantly differed with each other with respect to the variable Play (Pearson Chi-Square of 0.021). Business Females were significantly less likely to allow a child to Play than Art Males. No other Major-Gender pairing showed significant correlations with the variable Play. H6 is supported.

Discussion

Earlier research found that while all samples exhibited overwhelming non-acceptance of this breast-feeding doll for 5-year-olds, contrary to expectations, Arkansas disliked the doll less than California. Males accepted the doll more than Females in both Art and Business. Male Art students proved to be the most tolerant towards the doll and Female Business students the most intolerant [Brice et. al., 2016].

These results clearly showed that the Bebe Gloton doll product is culture-bound, both directly and through linking acceptance with gender, major, and major/gender combinations. It is reasonable that product acceptance and non-acceptance should be related to only one or two cultural constructs, as the seven cultural constructs examined here have little correlation. Market reaction related to one construct should not translate to other constructs. The fact that Power Distance was the only cultural construct significantly correlated with the variable Play supports the idea that it is of special importance. Other research has pinpointed Power Distance as having distinctive importance in differentiating family firm culture from non-family [Brice, Richardson, 2009]. In addition, H4 of this study supports the idea that Power Distance is the primary cultural construct that differs among the four Major/Gender groups.

High Power Distance is considered to be a socially conservative value. The association of choice of Play with higher Power Distance confirms earlier results based on demographics [Brice et. al., 2016]. Despite early assumptions, this research supports the idea that a breast-feeding doll product cannot be considered to be a progressive product and seems to indicate that the new breast-feeding social movement is not really accepted even by American progressives. Speculation might center on whether breast-feeding could be perceived as contrary to front-line progressive goals involving male-female career equality.

These results show that specific cultural constructs also correlate with College Major, Gender, and especially specific Major/Gender combinations. College major is possibly self-selected by individuals in accordance with their values and beliefs. This study goes beyond raw gender disparities between majors and reveals disparities in self-selected cultural values and beliefs within and between gender and major combinations. In particular, socially conservative values positively correlated with Business. More importantly, specific Major/
Gender pairings differed from each other with respect to six of the seven cultural constructs tested here. Simply comparing gender would not bring out the essential differences and clearly Gender and Major cannot be considered fully adequate variables on their own.

The question about whether this doll truly represents progressive or conservative values is partially answered by this study. Additional questions that arise include: is it because a breast-feeding doll touches strongly on Female issues that Females reject the doll more than Males? Perhaps the line between progressive and conservative is not linear. Social commentators often claim that “progressives”, being further from the mainstream, are more intolerant than centrist “liberals”. Are Art Males somewhat more “liberal” than Business students AND less “progressive” than Female Art students; and therefore more tolerant than Females at either end of the progressive/conservative spectrum?

This study supports the idea that this product might have a strong U.S. niche market of about ten percent, if those accepting the product have as much passion about it as the non-acceptors [Brice et. al., 2016]. Respondents were also aware that this product was Europe-based and Sweeney [1993] found that a culture gap may exist that tends to make this product seem ‘unhealthy’ and foreign. The disparities illustrated here show that there does not exist one national culture and that an imported product that appears to be rejected, has the potential to become successful if the market is segmented to find an accepting subculture [Thelen et. al., 2003]. Research shows that the Spanish importer of this product can overcome its basic lack of U.S. market knowledge by incrementally expanding its involvement in the U.S. [Johanson, Vahlne, 1977].

**Implications and Limitations of this Research**

This study supports the idea that, at least in a large masculine country like the U.S., the proper units of market investigation should consider region and gender/occupation groups rather than simply just gender or occupation. Assumptions cannot be made about the political or class identity of a product without both quantitative and qualitative research.

Cultural taboos can strongly and negatively affect product acceptance but a careful analysis of cultural segments can reveal strongly positive niche markets within the overall negative social culture. Even when a foreign product faces an overwhelming negative cultural response, some acceptance can be found in cultural subgroups. This acceptance may be as passionate as disapproval when that disapproval is based mostly on cultural attitudes. This is so because cultural attitudes vary widely between subcultures, even within the same nation. The implication is that products which are culture-laden need to be matched with the appropriate subcultures.

The major limitation of this research is that it is confined to only two regions and a single country. Further research is necessary to more precisely identify what cultural
attributes make a difference in terms of gender/occupation, ethnicity, age and other segments in terms of product taboos. Similar groups should be surveyed in additional regions within the U.S. This research should be replicated in Spain where this product originates and in Sweden, which displays more welfare-oriented values than does the U.S. [Hofstede, 1980, 2001], as well as in other countries shown to strongly differ from the U.S.

Notes

1 Author’s e-mail address: wbrice@csudh.edu
2 Author’s e-mail address: echu@csudh.edu
3 Author’s e-mail address: anastasiya.brice@yahoo.com

References


Appendix

Part One

Bebe Gloton Doll Survey

The purpose of this survey is to determine market reaction to a new doll imported from Spain. This survey is completely anonymous.

A Spanish toymaker has developed a breast-feeding doll that comes with a special halter top its young “mothers” wear as they pretend to breast-feed their “babies”. The halter top has daisies containing electronic sensors that cover the little girls’ breasts.

The doll, called Bebe Gloton which translates as “gluttonous baby”, makes sucking noises when its mouth is in proximity to the halter top sensors as it “feeds”. Like many other dolls, Bebe Gloton can cry, signaling she wants more milk and can be burped in the same way a real baby would be.

1. What is your academic major? _____________________________________

2. (circle one) Are you: (a) male or (b) female

3. (circle one answer) How old are you?
   - Under 18
   - 18–24
   - 25–29
   - 30–34
   - 35–39
   - 40–49
   - 50–59
   - 60 or over

4. What do you consider to be your primary culture (circle BEST answer and complete)?
   A. Not from America (please name country of origin ______________________)
   B. African-American (what is your home state? __________________________)
   C. Asian-American (what is your home state? ___________________________)
   D. European-American (what is your home state? ________________________)
   E. Hispanic-American (what is your home state? _________________________)
   F. Native-American (what culture and home state? ________________________)
   G. other American (what culture and home state? ________________________)

5. (circle one) Is there anything immoral about this doll if it is for a 5 year old girl?
   Yes  No  Not Sure
   If Yes, the moral issue is most related to (circle one):
   A. promotes early sexuality
   B. could involve the uncovering of breasts
   C. breast-feeding is not something to be encouraged
   D. other (please state) __________________________________

6. (circle one) If you had a 5 year old daughter, would you let her play with this doll?
   Yes  No  Not Sure

7. (circle one) If you had a 5 year old daughter, would you buy this doll for her?
   Yes  No  Not Sure
Part Two

Please think of an ideal job, disregarding your present job if you have one. In choosing an ideal job, how important would it be to you to...

(please circle one answer in each line across):  

1. have a good working relationship with your direct superior?  
   - of very little or no importance  
   - of little importance  
   - of moderate importance  
   - very important  
   - of utmost importance  

2. work with people who cooperate well with each other?  
   - of very little or no importance  
   - of little importance  
   - of moderate importance  
   - very important  
   - of utmost importance  

3. be consulted by your direct superior in his/her decisions?  
   - of very little or no importance  
   - of little importance  
   - of moderate importance  
   - very important  
   - of utmost importance  

4. have an opportunity for advancement to higher level jobs?  
   - of very little or no importance  
   - of little importance  
   - of moderate importance  
   - very important  
   - of utmost importance  

5. How frequently, in your experience, are subordinates afraid to express disagreement with their superiors?  
   - very frequently  
   - frequently  
   - sometimes  
   - seldom  

To what extent do you agree or disagree with each of the following statements?  
(please circle one answer in each line across):  

6. Most people can be trusted.      
   - strongly disagree  
   - disagree  
   - no opinion  
   - agree  
   - strongly agree

7. An organization structure in which subordinates have two bosses should be avoided at all costs.  
   - strongly disagree  
   - disagree  
   - no opinion  
   - agree  
   - strongly agree

8. When people fail in life it is often their own fault.  
   - strongly disagree  
   - disagree  
   - no opinion  
   - agree  
   - strongly agree

9. Religious faith contributes to good mental health.  
   - strongly disagree  
   - disagree  
   - no opinion  
   - agree  
   - strongly agree

10. Caution helps avoid mistakes.  
    - strongly disagree  
    - disagree  
    - no opinion  
    - agree  
    - strongly agree

11. Good luck follows if one survives a disaster.  
    - strongly disagree  
    - disagree  
    - no opinion  
    - agree  
    - strongly agree

12. Human behaviour changes with the social context.  
    - strongly disagree  
    - disagree  
    - no opinion  
    - agree  
    - strongly agree

13. Religion makes people escape from reality.  
    - strongly disagree  
    - disagree  
    - no opinion  
    - agree  
    - strongly agree

14. People may have opposite behaviours on different occasions.  
    - strongly disagree  
    - disagree  
    - no opinion  
    - agree  
    - strongly agree
<table>
<thead>
<tr>
<th></th>
<th>Statement</th>
<th>Levels</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>Fate determines one's successes and failures.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Religious people are more likely to maintain moral standards.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Ghosts or spirits are people's fantasy.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>There is a supreme being controlling the universe.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>One who does not know how to plan his or her future will eventually fail.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Art is a compensation for the dismal features of everyday life.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Knowledge is necessary for success.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Young people are impulsive and unreliable.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>It is rare to see a happy ending in real life.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>Individual characteristics, such as appearance and birthday, affect one's fate.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Adversity can be overcome by effort.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Every problem has a solution.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>One has to deal with matters according to the specific circumstances.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Competition brings about progress.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>There is usually only one way to solve a problem.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Tranquillity and serenity lead to a contented existence.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Most disasters can be predicted.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Old people are usually stubborn and biased.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Person's talents are inborn.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>One's behaviours may be contrary to his or her true feelings.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>There are certain ways to help us improve our luck and avoid unlucky things.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>One will succeed if he/she really tries.</td>
<td>1 2 3 4 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
37. Failure is the beginning of success.
38. Religious beliefs lead to unscientific thinking.
39. One feels safer in the world through a belief in a supreme being.
40. Current losses are not necessarily bad for one's long-term future.
41. Power and status make people arrogant.
42. Powerful people tend to exploit others.
43. People will stop working hard after they obtain a comfortable life.
44. Social institutions in society are biased towards the rich.
45. Belief in religion helps one understand the meaning of life.
46. After death, one exists in another form.
47. Kind-hearted people are easily bullied.
49. People deeply in love are usually blind.
50. Kind-hearted people usually suffer losses.
51. Caring about the larger affairs in society only brings trouble for yourself.
52. There are many ways for people to predict what will happen in the future.
53. Hard working people will achieve more in the end.
Exploring Links Between Engaging Customers in Value Co-Creation and Product Innovativeness

Abstract

The study examines the hypothesis that firms engaging customers in value co-creation tend to display more innovativeness. As such, it is one of the few quantitative studies on the link between these two concepts. Customer engagement in value co-creation was operationalized as a multiple scale following the DART framework by Prahalad and Ramaswamy. The DART acronym denotes four salient dimensions of enabling co-creation: Dialog, Access, Risk and Transparency. The applied innovativeness metric was revenue share from new and modified products. Data were collected from 432 managers of manufacturing and service SMEs. Statistical data analysis methods included EFA, CFA and multiple regression modeling.

The major finding is the existence of a significant positive effect between engaging customers in value co-creation and innovativeness. In particular, certain DART dimensions, such as Dialog, elements of Access and Risk, coincided with increased levels of innovativeness. Among the study’s limitations, two are particularly pertinent. First, different conceptualizations of customer engagement in value co-creation could yield different results in terms of effect magnitudes, although the authors believe that the direction of relationships should remain the same. Second, the research considered
customer engagement from the perspective of managers, which could induce bias. Hence, it may be worthwhile to examine how customers evaluate their own engagement.

In terms of practical application, to enhance innovativeness, firms should intensify their efforts to engage customers in day-to-day operations. However, not all aspects of co-creation provide equal benefits – it appears that more involved actions on the part of the company are needed to produce noticeable positive effects.

From a theoretical viewpoint, the findings empirically validate the business relevance of engaging customers in value co-creation. Unlike many other studies of the co-creation stream, this paper relies on a large, representative sample of manufacturing and service firms.

**Keywords**: value co-creation, DART model, service dominant logic, innovation, SMEs
**JEL**: L14, L15, L23, O32

**Introduction**

Our current knowledge driven economy is characterized by intense competitive pressures and the growing power of consumers, who have become more empowered by the ubiquitous presence of the Internet and related information technologies. In this setting, two themes are often cited as promising sources of lasting business success: engaging customers through value co-creation and perpetual innovation.

Co-creation with customers has become an increasingly important topic in the management and marketing literature concerning Service-Dominant Logic (SDL). An extensive literature review on value co-creation by Ranjan and Read [2014] identifies two types of research on this topic depending on the assumed understanding of value co-creation. It can be occurring either (1) in the process of experiencing the offer by the customer, or (2) when customers are engaged in value proposition development.

Accordingly, there are two manifestations of value co-creation. One is observed during the process of consuming/using a product by a customer (value-in-use), while the other occurs when firms and its customers work jointly on developing a value proposition (referred to by many authors as co-production). It should be noted that we purposefully avoid using the term co-production as somewhat misleading. Vargo and Lusch posit that the customer is always a co-producer while consuming the product [Vargo, Lusch, 2006, p. 18], obviously having in mind value-in-use, while Ranjan and Read [2014] define the term as direct or indirect “co-working with customers” or participation in the product/service design process, which is perhaps more appropriately referred to as value proposition co-creation.

Many authors suggest that companies can expect to gain multiple benefits from increased customer engagement in value co-creation, including strengthening brands, rising customer loyalty, attracting new customers [Piligrimiene et al., 2015; Saarijarvi
et al., 2013], as well as reducing various expenses to offer more differentiated products at lower prices [Hippel von, 2005; Arakji, Lang, 2007]. Co-creation can also enhance innovation through sharing, aggregating and recombining knowledge on the Internet [Brynjolfsson, McAfee, 2014], or by effective participation of customers in the designing and manufacturing of new products [Payne et al., 2008; Chen et al., 2012].

Some authors, however, have doubted whether a balanced, harmonious relationship between a firm and consumer is commonplace, and – more importantly – if co-creation consistently yields satisfactory business outcomes. Admittedly, in the new Internet-driven economy control over co-creation is visibly shifting to consumers, who can significantly influence brand value perceptions by mobilizing a powerful force – customer communities – to share their insights and experiences on products, brands and suppliers [Fisher, Smith, 2011]. This new consumer power can detrimentally affect “the profitability of conventional producer strategies that are based on pushing product designs that serve large segments of consumers while ignoring the service of more nuanced consumer preferences” [Lang et al., 2015].

With the caveat that value co-creation can create both threats as well as opportunities, we focus here on how engaging customers in value co-creation by Polish manufacturing and service SMEs affects their innovativeness.

The paper is structured as follows. The next section summarizes the relevant value co-creation and innovation literature. Research methods employed in data collection and statistical analysis are then discussed, and survey findings presented. The paper’s conclusions come next, along with suggestions for further research.

Literature Review

Evolving Nature of the Concept of Value

Use of the term “value” can be traced back to works by Plato and Aristotle but, as a concept, its meaning has changed throughout the years in various disciplines. Gummerus [2013] notes the common criticism that “value is one of the most misused terms” [Leszinski, Marn, 1997], and value research remains “an area of continuing ambiguity” [Woodall, 2003; Sánchez Fernández, Iniesta-Bonillo, 2007]. Our current, rapidly changing business environment has facilitated new approaches to value and how it is created in the management and marketing literature, along with efforts to better order and understand multiple approaches to value and their contexts. To that end, Ng and Smith [2012] presented a comprehensive literature review on value, contrasting traditional and modern approaches. Within the traditional stream, they identified five distinct value perspectives; that is, (1) utility, (2) economic worth, (3) perceived satisfaction, (4) net benefit and (5) a means towards a goal. Regardless of their differences, all of these views perceive
value as an intrinsic quality of an object, with passive and subjective customers that do not actively participate in value creation. Ng and Smith point out that this traditional value perception dates back to Adam Smith [1776], who distinguished between “value-in-use”, or the utility of some specific item, and “value-in-exchange”, representing the price commanded in the market. Smith also originated the concept of value embedded and distributed in tangible goods. Starting in the 1990s a new perspective on value began to gain ground [Holbrook, 1994, 1999, 2006]. This modern approach considers value as formed through the customer’s personal experience.

**Value in Service Dominant Logic**

The experience-based concept of value stands in opposition to this traditional interpretation in that – among other things – it postulates a different role for customers, who become value creators and equal partners to companies. Here, value is not embedded in a manufacturing or distribution process but is instead derived from product use and interactions between companies and customers [Vargo, Lusch, 2006]. The recent interest in experiential value co-creation observed in marketing literature was triggered by Vargo and Lusch [2004], who authored the foundations of Service-Dominant Logic (SDL). Similar to Holbrook [1987], Vargo, Lusch and Morgan [2006] claimed that customers create value by engaging in the process of product acquisition, usage and disposal. According to SDL proponents, companies cannot offer value as such, but rather value propositions [Frow, Payne, 2008] that become realized value when customers experience them in a unique context at a given time. This notion is reflected in how SDL looks at physical goods, which it assumes are akin to services because customers really buy the ensuing service that a tangible good provides when used [Vargo, Lusch, 2004]. Accordingly, SDL treats all offerings as services, which require consumption or use by a customer to yield value. One consequence of this approach, proffered by Lusch and Webster Jr. [2011], is a shift in marketing focus from sales and profit optimization to supporting customers in their value co-creation activities. This outlook is consistent with Gummesson’s opinion that “value creation is only possible when a good or service is consumed. An unsold good has no value, and a service provider without customers cannot produce anything”. [2002].

Though SDL has its critics (for example Campbell et al. [2013]), it has encouraged many academics to leave the constraints of the traditional marketing domain and contribute to the emerging paradigm [Li, Petrick, 2007; Warnaby, 2009; Löbler, 2011; Brodie et al., 2011; Laczniak, Santos, 2011; Karpen et al., 2012; Edvardson, Tronvoll, 2013; Edvardson et al., 2013]. This lively academic interest suggests the importance of further research in this area.

**Engagement in Co-Creation**

In this paper, in keeping with SDL [Vargo, Lusch, 2008], we understand co-creation as customer participation in developing a value proposition, named by Gustafsson et al.
[2012] as “innovation co-creation” or “co-creation for others”, and experience co-creation (“co-creation for use”) while realizing the value embedded in the value proposition through consumption or use. Specifically, the focus is on customer engagement in product innovations – setting up the specifics of the offering, and interacting with the company by sharing experiences. As such, the key to true co-creation is the notion of engagement, which indicates the depth of customer involvement. Among many relatively similar meanings of the term in marketing literature, we chose the one by Vivek et al. [2012], which is the most consistent with our research framework. Accordingly, consumer engagement is “the intensity of an individual’s participation and connection with organization’s offerings and activities initiated by either the customer or the organization” [Vivek et al., 2012].

Regardless of its role in SDL, customer engagement has been a popular research topic. Indeed, Marketing Science Institute listed it as a key research priority in the years 2010–2012 [Brodie et al., 2011]. Numerous authors perceived customer engagement as a strategic factor to enhance business performance directly [Voyles, 2007; Neff, 2007] or indirectly through supporting product innovation [Hoyer et al., 2010]. Kumar et al. [2010] pointed to a variety of ways customer engagement can support firms. Also, the concept of engagement was given a prominent role by authors studying customer experience [Prahalad, Ramaswamy, 2000].

In this study we considered the impact of customer co-creation from the perspective of a company’s efforts to engage customers, assuming that the more favorable environment is offered to customers, the more they are willing to get engaged. We did not investigate customers’ perceptions of being engaged by companies. Instead we study company efforts to practice co-creation with customers. Our literature review suggests that the DART model [Prahalad, Ramaswamy, 2004a] was arguably the most complete system, which considers co-creation in terms of a firm’s core activities, but also includes various supportive functions that could impact customer experiences. The DART model is described next.

The DART Model

Consistent with the co-creation mindset, Prahalad and Ramaswamy proposed their DART model as an aid for companies to enhance the customer role in the value creation and innovation processes [Prahalad, Ramaswamy, 2004a]. Their focus, arguably, is fostering conditions for building an effective communications platform with customers, which is an essential prerequisite for intense customer involvement with a company. Practically, the merit of DART lies in indicating the range of capabilities necessary for successful co-creation. In doing so, it specifies four main building blocks or competence groups that companies need to develop, and the “DART” acronym refers to those four components:

1. **Dialog** representing ongoing, unrestricted communication between a company and consumers performed on equal terms.
2. **Access** comprises tools and procedures that facilitate co-creation and increase freedom of choice for customers.

3. **Risk assessment** are measures that allow customers to fully evaluate the risk involved in accepting a value proposition.

4. **Transparency** is the extent to which a firm has managed to mitigate information asymmetry in relations between the company and its customers.

It should be noted that, when Prahalad and Ramaswamy first proposed this framework, they included so-called “dimensions of choice” to complement the DART model [2004a] by representing additional factors that can shape the overall co-creation experience by accounting for various modalities in product offerings. These modalities, or options, concerned customers’ ability to choose the most suitable distribution channels, communication variants, pricing, payment and financing methods. To keep our co-creation model simple and in agreement with several other authors [Albinsson et al., 2011, 2016; Ruso Spena et al., 2012], we opted against distinguishing dimensions of choice as a separate taxonomy, and instead to embed them within the Access component of DART. This extended the meaning of Access to include additional personalization options enhancing individual customer experience. The lack of separate dimensions of choice can also be found in later works by Prahalad, Ktishnan and Ramaswamy [Prahalad, Krishanan, 2008; Ramaswamy, 2008], which seems to further validate this decision.

Despite other attempts at conceptualizing and operationalizing co-creation [e.g. by Grönroos, 2009: Yi, Gong, 2013; Ranjan, Read, 2014] DART seems to be the most complete system. Unfortunately, previous attempts to test it empirically relied mostly on qualitative evidence.

### Customers as a Source of Innovation

Many insights about the role of customers in innovative firm processes are found in the works of Eric von Hippel [1976, 1977, 1986, 1988, 1994, 1998, 2001, 2005], who focused on user-centered innovation. According to him, due to information asymmetry “users and manufacturers know different things” [2005]. While users are acquainted with context-of-use information, specialized manufacturers have generic solution information. This is why users, led by their personal experiences, can suggest novel product functionalities, while manufacturers tend to offer improvements to better serve well-known needs. While using products, customers form opinions on their utilities and shortcomings. They also develop ideas for entirely new products or modified versions. Consequently, inviting customers to share their knowledge – before a competitor does so – is an important managerial task.

To increase competitiveness, firms tend to acquire valuable knowledge from multiple external sources (including customers) and combine it with internal knowledge resources. In this way, customer engagement in value proposition development is consistent with
the open innovation model, which is advanced by its proponents as an alternative to the vertical integration of the R&D function [Chesbrough, 2003]. Henry Chesbrough defined open innovation as “the use of purposive inflows and outflows of knowledge to accelerate internal innovation and expand the markets for the external use of innovation” [Chesbrough et al., 2006]. Dialog with customers as one source of knowledge may bring multiple benefits to the company, not the least of which are decreasing failure rates of new products due to a better understanding of the market and customer needs [Hippel von, 2005].

Works by other authors, Ranjan and Read [2014] indicate that customer engagement in developing a value proposition (co-production) requires collaboration [Lusch et al., 2007], dialog [Aarikka-Stenroos, Jakkola, 2012; Grönroos, 2012], interactivity, deep engagement and a willingness to share knowledge [Prahalad, Ramaswamy, 2004c]. Establishing and maintaining such collaboration and dialogue requires the involved parties to exhibit trust and openness. DART incorporates this basic willingness of customers to cooperate and share knowledge in an interactive, engaged dialogue.

Fisher and Smith point out that cooperation with customers may be fraught with difficulty. Many consumers are reluctant to have a close dialog with companies. Rather, they prefer to communicate with other peers. According to the authors, this is when “… companies need to turn into listeners and find ways to effectively eavesdrop on consumer conversations…” [2011], both on-line and off-line.

Previous Research on Co-Creation and Innovativeness

In previous studies, the relationship between co-creation and innovativeness was primarily concerned with the consequences of involving customers directly in companies’ innovation processes. The reported results were usually positive, the share of innovative products was higher and new products failure rates lower in firms that sourced novel ideas as well as other insights from their customers (an extensive review of works in this area can be found in Bogers et al. [2010]). To the best of our knowledge, the experience co-creation has not yet been adequately investigated in the innovation context on a large, representative sample of service and manufacturing companies. One of the few published works that did try to quantitatively identify links between experience co-creation and innovativeness used a statistical analysis of keywords on the web pages of a purposeful sample of 273 companies [Tanev et al., 2011]. The reliance on keywords, though a legitimate research method, has such limitations as omitting actual metrics of innovation activities of firms, and the non-representative sample did not permit generalization to a larger number of businesses.

Another study exploring links between co-creation and innovative performance considered the impact of customer dialog on innovativeness of Taiwanese IT-service
companies. This research surveyed 149 managers and found that firms more involved in dialog had better innovative outcomes [Hsieh, Hsieh, 2015]. Two major limitations of the study, however, included a narrowly defined industry focus, and establishing the level of innovative performance on the basis of Likert-scale items (the same method as for the other elements of the model), which might bias the results due to common measurement variance.

The brief literature overview outlined above indicates a knowledge gap, which is made more apparent by the presence of conceptual papers hypothesizing positive effects and attempting to explain possible mechanisms of the impact of experience co-creation on innovativeness [Tanev et al., 2009; Tanev, Frederiksen, 2014]. Addressing this knowledge deficiency, we use DART to reflect a wide range of co-creation practices.

With regard to innovation activity outcomes, our focus was on broadly considered product innovations, which comprise any modification to existing product or service structure, functional features, aesthetics, utility, as well as the introduction of entirely new offerings. In particular, as a measure of innovative success we used revenues from new and modified products as a percentage of sales for the year (2013) preceding data collection. This and similar metrics were employed in several other studies that measured the performance of innovation processes [e.g. Zeng et al., 2010; Ebersberger et al., 2012].

The growing prominence of co-creation practices in business and continuing, critical role of innovation to sustain competitive success makes it important to know how those two concepts interact. For that purpose, we argue (from the literature and our own observations) that a stronger involvement in activities enabling customer engagement in value co-creation is associated with higher levels of innovativeness. We set out to verify this hypothesis in the remainder of the paper by first discussing employed research methods, reporting the findings, and presenting our conclusions.

**Research Method**

**Sample Characteristics**

Data for the study were collected in July and August 2014 through CATI interviews and encompassed 432 cases for a response rate of approximately 30%. The interviewed managers represented small and medium Polish food manufacturing enterprises \( (n = 206) \), as well as hospitality, tourism (travel agents) and catering services \( (n = 226) \). To ensure representativeness, the sample was drawn from a database comprising virtually every manufacturing and service company in Poland, maintained by a research agency that was hired to administer the survey.

In the study, due to different interaction patterns with customers, a distinction was made between two groups of firms: manufacturers and service providers. As a rule, service personnel can directly observe customers using services and propose innovations based
on these observations. Manufacturers, however, typically lack this first-hand contact and need to purposefully inquire customers about their experiences with products, and provide them with a convenient and effective means for sharing those opinions. This distinction sets up an interesting comparison on the variables of value co-creation and innovativeness, and provides a partial basis to assess whether the patterns analyzed here are broadly applicable or rather specific to the manufacturing and service industries.

DART Measurement Scale

The central part of the research instrument was a multi-item Likert scale for establishing the extent of co-creation involvement according to the DART framework.

<table>
<thead>
<tr>
<th>DART component 1: DIALOG</th>
</tr>
</thead>
<tbody>
<tr>
<td>D 1: We systematically engage in dialog with consumers of our products/services.</td>
</tr>
<tr>
<td>D 2: We use special means to encourage consumers to have dialog with us.</td>
</tr>
<tr>
<td>D 3: One objective of our dialog with customers is generating ideas for new and modified products/services.</td>
</tr>
<tr>
<td>D 4: Dialog with consumers enables us to learn about their experiences with our products/services.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DART component 2: ACCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 1: Consumers can freely choose the time of product delivery/service provision.</td>
</tr>
<tr>
<td>A 2: Consumers can decide about certain elements of our products/services that influence their utility and/or the way they look.</td>
</tr>
<tr>
<td>A 3: Consumers can always choose their preferred payment method.</td>
</tr>
<tr>
<td>A 4: Consumers can always choose their preferred method of communicating with us.</td>
</tr>
<tr>
<td>A 5: Consumers can readily learn the specifics of our offer.</td>
</tr>
<tr>
<td>A 6: Information about our offer is easily available for consumers on our web site.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DART component 3: RISK ASSESSMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>R 1: Consumers can entirely consciously make their purchasing decisions because we inform them fully about the benefits of our products/services.</td>
</tr>
<tr>
<td>R 2: Consumers can entirely consciously make their purchasing decisions because we inform them fully about the risks from using our products/services.</td>
</tr>
<tr>
<td>R 3: We discourage from buying those consumers for whom we believe our products/services are not appropriate.</td>
</tr>
<tr>
<td>R 4: We encourage consumers to learn detailed information about using our products/services.</td>
</tr>
<tr>
<td>R 5: We modify user manuals and/or other information based on negative feedback from consumers about their product/service experiences.</td>
</tr>
</tbody>
</table>
In developing the statements for the scale we were guided by several earlier works, in particular Mazur and Zaborek [2014], Albinsson et al. [2011, 2016], Russo Spena et al. [2012], Prahalad and Ramaswamy [2000, 2004a, 2004b, 2004c]. Specific reasons for the content of the listed items are discussed below.

**Dialog Indicators**

In writing about dialog, Prahalad and Ramaswamy stressed its systematic, active and engaging character. This view was reflected in items D 1 and D 2. One objective of communicating through dialog is for both parties to learn about how to facilitate – among other things – better cooperation. It would not be valid to ask managers for their impressions of what insights their customers gained from dialog. But assuming a symmetrical, two-way relationship – which is implied by the very term dialog – an adequate proxy of dialog outcomes is obtained by asking managers what they learned from customers. This is addressed by D 3 and D 4. Also, it is worth noting that the word “dialog” is commonly understood in Polish (the language of the interviews) as a special kind of very close, fair and symmetrical communication, which accurately reflects the co-creation meaning of the word.

**Access indicators**

Consistent with Prahalad and Ramaswamy, who posit that access starts with information and tools [2004a], here Access is about letting consumers flexibly determine the preferred, personalized shape of an offer. This approach is equivalent to what Russo Spena et al. adopted in their research; that is: “Access covers how interaction empowers customer access to knowledge, tools, information and experience” [2012]. As such, to account for various modalities of choice, our questionnaire focused on: product or service features (A 2), distribution systems (A 1), payment methods (A 3), and information (A 4, 5 and 6). Our stronger emphasis on the information component corresponds to Prahalad and Ramaswamy’s focus on knowledge and information sharing in DART.
Risk Assessment indicators

In the original DART specification [Prahalad, Ramaswamy, 2004a] Risk Assessment was perceived as a means to lower “the probability of harm to the customer”. It seems that this goal can be accomplished by enabling consumers to make informed choices, which will result in a positive experience. Thus, consumers ought to have the right to learn about pertinent benefits and dangers involved in using a product (R 1 and 2), and the company should take a proactive stance in sharing this knowledge with its consumers (R 3 and 4). A correct implementation of the Risk Assessment function requires listening to customer feedback and implementing changes accordingly (R 5).

Transparency indicators

In keeping with Prahalad and Ramaswamy, Risk Assessment and Transparency contribute jointly to fostering trust, which is a foundation of fruitful value co-creation practices. Thus, their advice to managers was “When in doubt disclose” [2004c]. Since the role of Transparency in DART is to highlight the need to end information asymmetry in firm-customer relations, our questionnaire pertains to the quality of information offered to customers (T 1 and 2) and a company’s commitment to solving the problem of information asymmetry (T 3 through 6).

To sum up the overview of scale items, it should be noted that the DART framework, despite being a way to conceptualize value co-creation, does not have a fully validated set of scale items for use in quantitative research. In addition, there is a debate on the underlying structure of the latent variables [Mazur, Zaborek, 2014]. In particular, in the face of extant empirical evidence it is unclear if the scope and configuration of DART is best represented by only four constructs – as originally proposed by Prahalad and Ramaswamy [2004a] – or it is a more complex composition with more hidden variables accounting for more detailed aspects of co-creation. A likely reason for this uncertainty is that DART was first developed through case studies, and most subsequent research was also of the qualitative nature. Qualitative methods are known to use different scale validation techniques compared to a quantitative approach – in essence, qualitative researchers do not consider covariance matrices, but have to rely instead on various pattern matching schemes, as suggested by Yin [2011]. Hence, a measurement framework working well in the qualitative context may prove deficient in surveys. For this reason, the first step in statistical analysis involved us investigating the underlying structure of the scale items with exploratory factor analysis.
Methods of Statistical Analysis

To test the hypothesis of a positive link between the involvement of firms in value co-creation with consumers and innovativeness, we used exploratory factor analysis (EFA), confirmatory factor analysis (CFA) and multiple linear regression.

The purpose of the maximum likelihood EFA with SPSS 22 was to determine if the theoretical structure of the four latent DART variables also held for our data. As it turned out, the analysis identified seven hidden variables that correspond to Dialog, Transparency, two aspects of Access, two aspects of Risk and one new factor that we named Responsiveness. Responsiveness in communication between a firm and consumers means that there are no delays in the information flow, the contents of web pages and other media are up to date, and consumers can count on immediate replies and reactions. Consequently, in the second step, we relied on this latent variable structure to build and validate the measurement model via CFA, with the AMOS 22 software (see Table 2 for results). Once the measurement model yielded acceptable fit metrics the factor scores were saved as new variables and used as predictors in the multiple regression model, which concluded our statistical analysis.

An explanation is in order as to why we decided against building a full structural equation model (SEM) to verify the hypothesis. The main reason was that the graphical displays for the equivalent SEM would be too complex, and their adequate presentation too extensive for the scope of this paper. Moreover, in the context of this research once the CFA was completed, the SEM would not offer any substantive benefits over a multiple regression analysis. Specifically, one would have to present two models separately for manufacturing and service companies. Each would have to accommodate a large set of rather confusing covariance links between the seven first order factors. Also, the multiple regression made it easier to represent interactions between predictors, which we suspected could be present in the data.

Research Findings

As the starting point the CFA results will be reported. Here, the major issue was to see if manufacturing and services firms exhibited distinct differences in their respective DART measurement models. If not, then a single pooled measurement model could be used; otherwise two models would be needed. To address this issue a measurement invariance test was performed. In the test two models were statistically compared: the first with each group having independently estimated parameters (the so called unconstrained model assuming that the two types of firms are different) and the second having the same regression weights for both groups (i.e. the specification with constrained measurement
weights assuming that both groups of companies were described by the same DART model). A resulting significant chi-square value of 22.066 (p = 0.016) indicated that the firms could not be pooled together and treated as a single group and instead the two models should be estimated separately. The unconstrained model had a significantly better data fit than the constrained one, but this should not be taken to mean that each and every parameter was different across both models. In Table 2, the significance of specific pairwise differences was indicated whenever critical ratios fell beyond the range of −1.96 and 1.96 and, in these cases, regression weights were marked in bold.

### TABLE 2. CFA measurement model of DART framework for manufacturing and services firms

<table>
<thead>
<tr>
<th>Indicator content</th>
<th>Regression weights from the parent construct to the indicator</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>for manufacturers</td>
<td>for services</td>
</tr>
<tr>
<td><strong>Dialog</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Manufacturers: AVE = 0.62, Cronbach’s Alpha = 0.81, MSV = 0.18; Services: AVE = 0.56, Cronbach’s Alpha = 0.76, MSV = 0.48)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 1: We systematically engage in dialog with consumers of our products/services.</td>
<td>0.92</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>D 2: We use special means to encourage consumers to have dialog with us.</td>
<td>0.98</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>D 3: One objective of our dialog with customers is generating ideas for new and modified products/services.</td>
<td>0.59</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>D 4: Dialog with consumers enables us to learn about their experiences with our products/services.</td>
<td>0.58</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td><strong>Access 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Manufacturers: AVE = 0.54, Cronbach’s Alpha = 0.68, MSV = 0.14; Services: AVE = 0.60, Cronbach’s Alpha = 0.62, MSV = 0.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A 1: Consumers can freely choose the time of product delivery/service provision.</td>
<td>0.74</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>A 2: Consumers can decide about certain elements of our products/services that influence their utility and/or the way they look.</td>
<td>0.68</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td><strong>Access 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Manufacturers: AVE = 0.51, Cronbach’s Alpha = 0.65, MSV = 0.22; Services: AVE = 0.28, Cronbach’s Alpha = 0.61, MSV = 0.56)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A 3: Consumers can always choose their preferred payment method.</td>
<td>0.41</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>A 4: Consumers can always choose their preferred method of communicating with us.</td>
<td>0.76</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td>A 5: Consumers can readily learn the specifics of our offer.</td>
<td>0.88</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>Indicator content</td>
<td>Regression weights from the parent construct to the indicator for manufacturers</td>
<td>for services</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td><strong>Risk 1</strong> <em>(Manufacturers: AVE = 0.72, Cronbach’s Alpha = 0.71, MSV = 0.31; Services: AVE = 0.56, Cronbach’s Alpha = 0.64, MSV = 0.53)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 1: Consumers can entirely consciously make their purchasing decisions because we inform them fully about the benefits of our products/services.</td>
<td>0.69</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>R 2: Consumers can entirely consciously make their purchasing decisions because we inform them fully about the risks from using our products/services.</td>
<td>0.98</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td><strong>Risk 2</strong> <em>(Manufacturers: AVE = 0.52, Cronbach’s Alpha = 0.82, MSV = 0.31; Services: AVE = 0.61, Cronbach’s Alpha = 0.79, MSV = 0.35)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R 3: We discourage from buying those consumers for whom we believe our products/services are not appropriate.</td>
<td>0.74</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>R 4: We encourage consumers to learn detailed information about using our products/services.</td>
<td>0.77</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>R 5: We modify user manuals and/or other information based on negative feedback from consumers about their product/service experiences</td>
<td>0.64</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td><strong>Transparency</strong> <em>(Manufacturers: AVE = 0.44, Cronbach’s Alpha = 0.63, MSV = 0.25; Services: AVE = 0.28, Cronbach’s Alpha = 0.48, MSV = 0.56)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T 1: All information that we disseminate is reliable.</td>
<td>0.78</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>T 3: We follow an open information policy since we have nothing to hide.</td>
<td>0.88</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>T 5: We don’t try to hide when we are blamed for something; we address such charges openly.</td>
<td>0.48</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>T 6: If we happen to make mistakes, we admit to them publicly.</td>
<td>0.43</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td><strong>Responsiveness</strong> <em>(Manufacturers: AVE = 0.68, Cronbach’s Alpha = 0.82, MSV = 0.25; Services: AVE = 0.51, Cronbach’s Alpha = 0.75, MSV = 0.46)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A 6: Information about our offer are readily available for consumers on our web site.</td>
<td>0.89</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>T 2: Information published on our web site is up to date.</td>
<td>0.81</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>T 4: We immediately reply to questions from our current and potential customers.</td>
<td>0.77</td>
<td>0.49</td>
<td></td>
</tr>
</tbody>
</table>

**Overall fit metrics for the entire measurement model**: Chi-squared/df=1.344; GFI=0.908; AGFI=0.869; CFI=0.928; RMSEA=0.028; PCLOSE=0.876.

Source: own study.
According to the threshold levels given in Garson [2012], the general fit indices shown at the bottom of the table imply that the measurement model fits the data well, meaning that the covariance matrix computed from the model closely resembles the empirical covariance matrix. However, indicators obtained for individual constructs clearly point to some deficiencies in the model. Most notably, it is clear that DART is better suited for manufacturing companies than for service providers. Among manufacturers, each construct has a satisfactory level of average variance extracted from its indicators (AVE of more than 50%) and there are no issues with discriminant validity (AVE is always greater than MSV, standing for maximum shared variance). For service companies some constructs are rather difficult to interpret due to low AVE coefficients, and AVE being lower than MSV. Specifically, the factors Risk 2 and Transparency explain little variance in its indicators and are too similar to other factors. This finding is consistent with our previous research on DART [Mazur, Zaborek, 2014], where – while using a different sample – it was concluded that the DART framework works better for manufacturing companies, possibly because it was originally developed through a qualitative investigation of several manufacturing firms. Even though services firms are generally worse represented by the model, for factors other than the problematic Risk 2 and Transparency reliability and validity is at least adequate, and Dialog and Risk 1 have their measurement models equivalent to manufacturers’ (as indicated by insignificant differences between regression weights). This could imply that those two aspects of value co-creation are universal and valid regardless of the nature of business operations.

As a general note, the regression weights in Table 2 should not be interpreted as telling what indicators are more prevalent among which types of firms, nor that manufacturing firms are involved more in value co-creation. Rather, they show where the DART model is more accurate and reliable.

When interpreting the “split” factors labeled Risk 1 and 2, and Access 1 and 2 it appears that Risk 1 encapsulates more passive aspects of addressing the danger and inconvenience involved in purchasing wrong products or inadequate services, as it entails providing correct information to consumers. In contrast, Risk 2 is more proactive, since it requires specific actions targeting particular needs and circumstances of individual customers. Access 1 concerns practices that can provide substantial benefits to customers, such as home delivery or custom made products. Access 2 components involve more intangible characteristics of the offer, relying primarily on the availability and the efficient transfer of information. Hence, it indicates that dividing Risk and Access was grounded in substantive reasons, and was not merely driven by statistical criteria.

A better understanding of the components of the DART measurement model can be gained from Table 3 displaying Pearson correlation coefficients between pairs of latent variables.
TABLE 3. **Bivariate correlations between elements of the DART model for manufacturing and services**

<table>
<thead>
<tr>
<th>Pairs of correlated variables</th>
<th>Manufacturing</th>
<th>Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dialog &lt;--&gt; Access 1</td>
<td>0.254</td>
<td>0.252</td>
</tr>
<tr>
<td>Dialog &lt;--&gt; Access 2</td>
<td>0.399</td>
<td>0.415</td>
</tr>
<tr>
<td>Dialog &lt;--&gt; Risk 1</td>
<td>0.142</td>
<td>0.211</td>
</tr>
<tr>
<td>Dialog &lt;--&gt; Risk 2</td>
<td>0.089</td>
<td>0.165</td>
</tr>
<tr>
<td>Dialog &lt;--&gt; Transparency</td>
<td>0.072</td>
<td>0.188</td>
</tr>
<tr>
<td>Dialog &lt;--&gt; Responsiveness</td>
<td>0.069</td>
<td>0.089</td>
</tr>
<tr>
<td>Access 1 &lt;--&gt; Access 2</td>
<td>0.274</td>
<td>0.342</td>
</tr>
<tr>
<td>Access 1 &lt;--&gt; Risk 1</td>
<td>-0.029</td>
<td>-0.012</td>
</tr>
<tr>
<td>Access 1 &lt;--&gt; Risk 2</td>
<td>0.159</td>
<td>0.074</td>
</tr>
<tr>
<td>Access 1 &lt;--&gt; Transparency</td>
<td>0.055</td>
<td>0.136</td>
</tr>
<tr>
<td>Access 1 &lt;--&gt; Responsiveness</td>
<td>0.017</td>
<td>0.054</td>
</tr>
<tr>
<td>Access 2 &lt;--&gt; Risk 1</td>
<td>0.206</td>
<td>0.372</td>
</tr>
<tr>
<td>Access 2 &lt;--&gt; Risk 2</td>
<td>0.455</td>
<td>0.311</td>
</tr>
<tr>
<td>Access 2 &lt;--&gt; Transparency</td>
<td>0.305</td>
<td>0.427</td>
</tr>
<tr>
<td>Access 2 &lt;--&gt; Responsiveness</td>
<td>0.556</td>
<td>0.577</td>
</tr>
<tr>
<td>Risk 1 &lt;--&gt; Risk 2</td>
<td>0.523</td>
<td>0.571</td>
</tr>
<tr>
<td>Risk 1 &lt;--&gt; Transparency</td>
<td>0.192</td>
<td>0.378</td>
</tr>
<tr>
<td>Risk 1 &lt;--&gt; Internet use</td>
<td>0.094</td>
<td>0.16</td>
</tr>
<tr>
<td>Risk 2 &lt;--&gt; Transparency</td>
<td>0.051</td>
<td>0.14</td>
</tr>
<tr>
<td>Risk 2 &lt;--&gt; Responsiveness</td>
<td>0.123</td>
<td>0.102</td>
</tr>
<tr>
<td>Transparency &lt;--&gt; Responsiveness</td>
<td>0.537</td>
<td>0.622</td>
</tr>
</tbody>
</table>

*Source: own study.*

Correlations displayed in the table suggest that all estimated latent variables are parts of a single interconnected system. Also, correlational patterns for manufacturers and service providers are roughly similar, pointing to a broad equivalence of this aspect of the measurement model. One particular insight concerns the Responsiveness variable. This construct was introduced to accommodate part of unexpected correlational patterns that emerged in the structural equation analysis. Based on its associated Likert-scale items, it could be interpreted as the capacity of a company to keep its information flows up-to-date, which implies reacting quickly to consumer feedback and frequently amending the content of their web pages and other means of communication. From the correlation coefficients in Table 3 it seems that Responsiveness is not an intrinsic and exclusive part of any single DART element but it is related to all of them – either directly (Transparency and Access 2) or indirectly (the rest of constructs through links with Transparency and
Access 2). Statistically, the measurement model is more robust if the three manifestations of Responsiveness are brought together under a distinct latent variable. This has also a substantive appeal, since Responsiveness seems to be enabled by the use of Internet technologies and in the literature many aspects of value co-creation are believed to be facilitated by the rise of the internet. This obvious link of Responsiveness with the main driving force of value co-creation justifies – in our view – the inclusion of this variable in our model.

To facilitate further analysis, the regression weights of Table 2 were used to create factor scores for each firm in the sample, which resulted in 7 new variables representing the respective aspects of value co-creation. These new variables were entered into a regression equation as predictors. In addition, to control for two important characteristics of firms, the regression model was complemented by an economic sector variable (0 for manufacturing and 1 for services) and the size of the company measured by the number of employees. The final model specification step was to check for interactions between economic sector and other variables and input significant interaction terms into the equation. The resulting model is depicted in Table 4.

### Table 4. Parameter estimates of the multiple regression model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>B</th>
<th>Std. Error</th>
<th>t</th>
<th>p-values</th>
<th>95% Confidence Interval</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dialog</td>
<td>4.290</td>
<td>1.662</td>
<td>2.581</td>
<td>.010</td>
<td>1.022 – 7.558</td>
<td>.016</td>
</tr>
<tr>
<td>Access 1</td>
<td>3.802</td>
<td>1.719</td>
<td>2.211</td>
<td>.028</td>
<td>.422 – 7.182</td>
<td>.012</td>
</tr>
<tr>
<td>Risk 1</td>
<td>1.418</td>
<td>2.302</td>
<td>.616</td>
<td>.538</td>
<td>-3.107 – 5.943</td>
<td>.001</td>
</tr>
<tr>
<td>Risk 2</td>
<td>5.197</td>
<td>1.814</td>
<td>2.864</td>
<td>.004</td>
<td>1.630 – 8.763</td>
<td>.020</td>
</tr>
<tr>
<td>Transparency</td>
<td>-1.527</td>
<td>2.733</td>
<td>-559</td>
<td>.577</td>
<td>-6.899 – 3.845</td>
<td>.001</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>3.487</td>
<td>3.371</td>
<td>1.034</td>
<td>.302</td>
<td>-3.141 – 10.115</td>
<td>.003</td>
</tr>
<tr>
<td>No. of Employees</td>
<td>1.405</td>
<td>1.267</td>
<td>1.109</td>
<td>.268</td>
<td>-1.086 – 3.896</td>
<td>.003</td>
</tr>
</tbody>
</table>

Notes: Dependent Variable: Percentage of New and Modified Products in Total Revenues; R-squared for the model = 0.181.
Source: own study.

As indicated by the R-squared, the whole set of predictors accounted for 18.1% of the variance in the dependent variable. Further investigation of the model quality did not reveal any issues with multicollinearity, non-normal distribution of residuals or heteroscedasticity.
There were six significant parameters in the model: the intercept, Dialog, Access 1, Risk 2, Economic Sector and the interaction between Responsiveness and Economic Sector. Effect sizes of individual predictors can be determined from eta squared coefficients, which reflect the proportion of unique variance in the dependent variable explained by each predictor [Field, 2013]. The variable with seemingly the strongest effect on the Percentage of New and Modified Products was Economic Sector, which was a dichotomous attribute coded 0 for manufacturing and 1 for services. Here, with all other predictors held constant, the share of revenues from new and modified products generated by service providers was, on average, 14.96 percentage points higher than the share of manufacturers. The second most important predictor was Risk 2, closely followed by Dialog 1, and Access 1. For these aspects of value co-creation the b values were positive, meaning that increased levels of involvement, *ceteris paribus*, corresponded with higher levels of innovativeness. Relatively weakest was the effect of interaction speed, whereby increased Responsiveness was more characteristic of less innovative service companies (coded as 1 on the Economic Sector variable), while in manufacturing firms Responsiveness was not linked with any systematic differences in innovativeness. The intercept had a practical interpretation since all metric predictors (here, non-dichotomous) were standardized before entering into the model, and the only categorical variable – Economic Sector – could legitimately take the 0 value. As such, the intercept was the percentage of new and modified products in the revenues for a hypothetical manufacturing company with all metric predictors set to their means (the mean is 0 for a standardized variable). This number (24.89%) was very close to the average for all manufacturers (23.99%) and was quite distinct from the mean for service companies (38.95%).

To summarize, there was a significant positive effect between involvement in value co-creation operationalized by the DART model and innovativeness. However, not all aspects of VCC were associated with higher levels of innovation, which gives only partial support to the research hypothesis.

**Discussion**

The study findings are in general agreement with Gustafsson et al. [2012], who surveyed 334 managers to investigate communicative aspects of value co-creation in terms of frequency, direction, modality and content (it should be noted, though, that in contrast to our project the scales employed by these authors were more concerned with innovation co-creation than experience co-creation). They found that the frequency, direction and content of communication with customers corresponded to higher innovation success rates when innovations were incremental. For radical innovations, only the frequency of communication mattered. Even though we did not distinguish in our survey between
radical and incremental innovations, it seems certain that the vast majority of innovations in the investigated industries were of the latter kind.

Considering that dialog is “a form of reflective conversation that enables actors to alter managers’ mental models through conscious, critical exploration” [Jacobs, Heracleous, 2005] it is not surprising that its association with innovation was positive. After all, innovation – especially that of a strategic kind – calls for shifts in existing mental models. Intense dialog with customers was formerly found to be conducive to more intense and effective innovation practices (e.g. see a case study by Ayuso et al. [2006]), however this evidence was qualitative in nature, and so not directly comparable to the current research.

One previous quantitative study that found dialog with consumers supportive of better innovative performance was a survey of 149 managers of Taiwanese IT firms [Hsieh, Hsieh, 2015]. There the assumed concept of dialog (dubbed “dialogic co-creation”) encompassed three subconstructs labeled as “customer initiative”, “provider initiative” and “continuity”. The Likert-type indicators employed to estimate the three subconstructs suggest that the domain of the dialog construct also partially included the DART elements of Transparency and Access. The findings show that the impact of dialog co-creation is significant, positive and occurs through mediating variables of “company-customer relationship”, “knowledge valuation” and “customizing capacity”. Thus, the outcomes are similar to those of our study, except here the model was framed without mediating variables assuming only direct regression paths between DART and innovativeness. This consistency of findings, despite methodological differences, suggests that dialog with consumers is a real, beneficial process that transcends cultural and industrial boundaries.

Another significant predictor in the model (Access 1) involved highly interactive practices that engaged consumers in the process of determining features of their desired offerings. Such interactions, in addition to providing increased utility to customers, are information rich for firms, which naturally leads to a higher chance of acquiring useful insights for product innovation. This observation of stronger positive effects on the innovative function of deep involvement with consumers is consistent with previous case study research identifying multiple benefits of engaging customers through highly-interactive Web 2.0 tools, both in terms of innovation and day-to-day business operations [Martini et al., 2014; Martinez Garcia, 2013]. Notably, the two cited papers investigated food and beverage industry firms (the same as the manufacturers in our study), making it more likely that our findings are quantitative reflections of similar underlying mechanisms.

The Risk 2 variable represents the active involvement on the part of the company to ensure that its products are bought only by customers who can fully benefit from their features, sometimes at the cost of discouraging those customers for whom the products would not be suitable. Such an attitude and behavior promotes trust, facilitates dialog, and gives additional opportunity to glean insights from customers, which can inspire creative ideas for modifications and completely novel offerings.
Interestingly, in the present research service firms showed a mildly negative association between Responsiveness and revenues from innovative products. One possible explanation is that many service companies in hospitality, tourism and catering rely on daily face-to-face dialogue with customers. These direct interactions provide ample opportunities for dialog with consumers, making for richer communication than internet exchanges. The lack of a positive impact of Responsiveness on innovativeness may be true when only very simple communication tools are used. More advanced communication techniques frequently result in valuable innovations. One such example can be found in a multiple case study of seven hotels, where high Responsiveness from the use of social media and other deeply involving on-line applications helped boost customer satisfaction and generate ideas for product innovation [Shaw et al., 2011].

**Academic and Practical Implications**

From the academic perspective, our study provides further empirical evidence for the validity of the DART model, suggesting that its actual measurement structure could involve 7 instead of 4 hidden variables. Moreover, this conceptualization of co-creation appears to be more suitable for describing manufacturers than service providers but, in terms of its Dialog and Risk 2 components, both industries seem to show similar patterns of association. Evidence was also found of positive links between engaging customers in co-creation and innovativeness. As such, the obtained findings provide empirical support for one of the key foundational premises of SDL by showing that value co-creation could lead to business benefits derived from more and better ideas for product modifications and new concepts.

From a practical viewpoint, this study suggests that managers should consider increasing firm interactions with customers before, during and after product purchases. This may not only enhance customer experiences and product satisfaction, but also stimulate innovativeness. It is worth noting, though, that not all forms of co-creation provide equal benefits – it seems that only more active, involved efforts by a company can produce noticeably positive effects. These include conducting regular, two-way conversations with customers, offering them meaningful product configuration options and openly informing them of the risks and disadvantages of goods and services, even at the expense of losing those buyers who will find the offering inadequate to their needs. In the long run, however, this approach could increase a company’s reputation, bringing in new customers and strengthening the loyalty of the existing ones.
Limitations and Further Research

One limitation of this study is how we operationalized the value co-creation concept. Admittedly, the DART model is not perfect and a different approach could bring different results. However, considering that dialog, or its close equivalents, are the central elements of any co-creation concept, we are convinced that with alternative methods of operationalizing the possible differences will likely concern only the magnitude of the effect sizes and not the overall direction of relationships, which should remain positive.

The research design employed, even though it allowed us to estimate the general effect patterns, did not show precisely in which processes, e.g., technology and product development, manufacturing or commercialization, the insights from customers were used. Overcoming this limitation could be a valid objective of new research that may – for example – follow the framework outlined by Theyel [2012].

Also, the fact that the study was carried out in a single country and on a small set of industries (food, beverage, hotels and travel services) does not guarantee that the same patterns exist across other countries and industries. Finally, asking managers about their companies via quantitative interviews (though widely used in management science) may induce certain biases. It seems that several in-depth case studies on companies similar to those involved in the conducted survey could yield interesting, complementary insights about the underlying causal mechanisms linking co-creation with innovativeness. In addition, it would be worthwhile to attempt to replicate our research on samples drawn from different industries and countries. As the final suggestion for a follow-up study, it would be highly informative to run a survey with competing operationalizations of value co-creation to evaluate their validity and reliability, and attempt to develop a new, possibly more accurate, synthetic measurement approach.

Notes

1 Author’s e-mail address: pzaborek@gmail.com
2 Author’s e-mail address: jolanta.mazur@sgh.waw.pl
References


Identification of the Flow of Innovations in Tourism Related to Aesthetic Medicine

Abstract

Aesthetic medicine is, next to the wellness and spa, one of the most rapidly growing segments of health tourism. Its dynamic growth is closely linked to innovative offers (perhaps better described as “product innovation”). To date, however, there have been no scientific studies focusing on this market. Services in the field of aesthetic medicine are usually discussed descriptively as a subcategory of medical tourism, and innovation in this sector remains unexplored.

In this article we focus on innovation in aesthetic medicine as it pertains to health tourism, using the Delphi method to analyze innovation flow. Twelve experts from the fields of economics and management sciences with backgrounds in innovation and the economics of tourism (including health tourism) were invited to participate in the study. This group included also practitioners. The research was conducted in June and July 2015.
Our research addresses the current dearth of academic works on aesthetic medicine tourism, and lack of any relevant models of its flow and spread.

**Keywords:** aesthetic medicine tourism, aesthetic medicine, aesthetic medicine innovation, innovation flow channel  
**JEL:** O31 I15

## Introduction

As Europeans age, tourist services in aesthetic medicine are gaining importance. Aesthetic Medicine Tourism has emerged in response to changing social preferences related to improvement of mental and physical aesthetics associated with physical appearance. The success of the medicine tourism sector depends on innovative activities offered by this sector.

In the article we identify the flow of innovation in aesthetic medicine tourism sector, and identify providers of aesthetic medicine tourism services responsible for it. There is limited literature on this topic, and no modeling of its flow and spread. The presented research fills this knowledge gap.

## Development of Aesthetic Medicine Tourism

The limited literature that does exist on aesthetic medicine tourism places it within health tourism (combining aesthetic medicine and travel). Tourist services are well-defined in the literature, however the area of aesthetic medicine services requires further exploration. Aesthetic medicine is of recent vintage. Its roots are linked to: the creation of the French Society of Aesthetic Medicine (1973) and establishment of the Union International de Medicine Esthetique – UIME (1975) based in Paris. Aesthetic medicine in Poland dates back to the establishment of the Aesthetic Medicine Section at the Polish Medical Association in 1993, whose activities consisted of meetings, conferences and congresses. In 2002 the Graduate School of Aesthetic Medicine was opened, in 2007 the quarterly “Academy of Aesthetic and Anti-Aging Medicine” began publication, and in 2008 a certification program for doctors of aesthetic medicine was launched. In 2010 the Aesthetic Medicine Section became the Polish Society of Aesthetic Medicine and Anti-Aging which, by 2013, had more than 1000 physician members (Polish Association of Aesthetic Medicine).

Aesthetic Medicine deals with patient aesthetics, health, and psychophysical welfare [Ignaciuk, 2009, pp. 223–226]. Use of the adjective “aesthetic” denotes the use of medical procedures and products to target physical attractiveness (beauty), rather than disease
treatment. Individual medical specialties in the “aesthetics” segment often use techniques typical of other medical fields (e.g., drugs, procedures, nutrition, rehabilitation and physical therapy) but with the more narrow primary objective of improving physical attractiveness. Aesthetic medicine practitioners are physicians who use medical methods to meet this objective [Śpiewak, 2012, pp. 69–71]. Secondarily, aesthetic medicine also seeks to prevent, treat, correct, and provide rehabilitation for broader patient health issues [Ignaciuk, 2009, pp. 223–226].

Health tourism describes health and leisure needs, as well as the improvement of beauty and well-being [Lewandowska, 2007 pp. 174–190]. Since aesthetic medicine tourism is focused on removing external appearance defects and improving mental state, it is properly understood as a health tourism segment, because it motivates a person to travel [Gaworecki, 2003, p. 38].

The health tourism literature currently does not concern aesthetic medicine tourism, per se [Hunter-Jones, 2003, p. 170; Mika, Ptaszycka-Jackowska, 2007, p. 279; Lubowiecki-Vikuk, 2010, pp. 93–104; Prochorowicz, 2008, p. 133]. In our view, separate classification is warranted because aesthetic medicine tourism lies in between medical tourism, wellness and spa, and spa tourism, without replicating any one of them. For example, it includes improving body appearance through minimally invasive or non-invasive methods, without surgical intervention. This distinguishes aesthetic medicine tourism from other health tourism. The common element in all related segments is the combination of therapy and sightseeing.

**Innovativeness of Aesthetic Medicine Tourism**

Innovation is connected with the terms “idea”, “invention”, or “change”, “improvement”, “reform” [Innowacje…, 2005, pp. 65–66]. It is also referred as a specific tool for entrepreneurs that contributes to starting a new businesses, providing a new quality of services [Drucker, 1985, pp. 35–36]; or introducing a new product, process, system or device [Freeman, 1982, p. 7]. This applies to new applications of old solutions [Rogers, 1962, p. 143] and new solutions leading to developmental changes [Domanowska, 2006, p. 198]. Polish legislation defines innovation as: activities related to the preparation and launching of the production of new or improved materials, products, devices, services, processes or methods, intended to be introduced to the market or for other practical uses [Ustawa, 2005]. Innovation in tourism depends on creating new products and services, including those based on new social services and directions of travel [Damanpour, 2005, pp. 555–590]. It is important to point out that they are based on combining different types of social services and converting them into a commercial tourism product, often in response to changing social preferences.
The combination of aesthetic medicine treatments with tourist services (related to travel and stay) have led to the creation of an innovative tourism product that combines aesthetic medicine services with leisure tourism. It relies on new aesthetic medicine products and new preferences in the field of traditional tourism. The innovativeness of aesthetic medicine services is also illustrated by a rapidly growing pharmaceutical market as well as new pharmaceuticals and medical equipment. The merging of these sectors is also reflected in the mitigation of medical procedures and development of new therapeutic institutions, especially in the private sector. Since innovations in aesthetic medicine can influence the development of tourism, and tourist preferences can influence the search for medical products (and their institutionalization) the development of medicine may be a driving force for the tourism industry, and medical innovation can be a call for action for decision-makers in tourism.

The study of innovation in aesthetic medicine tourism should therefore recognize institutional linkages between tourism and aesthetic medicine. These links are responsible for the flow of information from one sector to another [Nootenboom, 2000]. When traditional boundaries and barriers of institutional linkages are exceeded, the creation of innovation occurs [Garcia-Altes, 2005, pp. 262–266].

Methodology of Research

The Delphi method enables the synthesis of essential knowledge (including implicit one) for the examined subject (Popper), and is used here to identify the directions of the flow of innovations in aesthetic medicine tourism.

Use of an expert panel and the Delphi method is based on four assumptions [Plummer, Armitage, 2007], which are:

- a group of participants (experts) selected based on their expertise;
- the process of multiple interactions (here double) through which expert opinions are exchanged and consensus achieved;
- participant feedback to facilitate interaction and reflection; and
- the opinions of experts that contribute to the solution of a selected problem.

The last step involves analyzing collected research materials using quantitative and qualitative methods [Loo, 2002, pp. 762–769].

The study was conducted in three stages. In the first stage, a team of contractors formulated the research theses from a theoretical analysis of the literature [Nazarko, 2013a, p. 7]. The study was performed using a CAWI questionnaire. The analysis involved 12 economic science experts with a background in innovation and the economics of tourism, including health tourism. Practitioners were also included in the expert team.
The second stage of the study focused on the performance of the first round of the Delphi survey and involved sending the CAWI survey to experts. After receiving completed forms, we performed an aggregated analysis of the results of the first survey round which was also sent to the experts.

In the third stage, experts once again received the research form and, in addition, a graph portraying the collective opinions of all experts participating in the first Delphi survey round. The experts could then change their opinion on a given subject, select new answers or maintain their original response [Rowe, Wright, 1999, pp. 353–375].

The study was conducted in the months of June and July 2015.

The structure of the analysis of the test results

The study identified:
- selected health tourism segments in which innovation would be analyzed;
- specific operators directly involved in the provision of aesthetic medicine tourism services, and their relative roles in this segment;
- entities forming the chain network of providers of aesthetic medicine tourism services that significantly impacted the flow of innovation; and
- the role of different participants in the innovation process.

The significance of service providers in developing aesthetic medicine tourism was performed using a scale of 0 to 5, where 0 means no significance in the provision of aesthetic medicine tourism services, while 5 means the greatest significance in the provision of the aesthetic medicine tourism services.

In each analyzed category the influence of service providers on innovation, the impact encouraging indicator (Ws) was designated (according to the formula) [Nazarko, 2013b, p. 106; Ejdys, 2013, p. 112; Dębkowska, 2013, p. 97]:

\[
W_s = \frac{n_{BD} \cdot 100 + n_D \cdot 80 + n_S \cdot 60 + n_N \cdot 40 + n_{BN} \cdot 20 + n_{BZ} \cdot 0}{n}
\]

where:
- \(n_{BD}\) – number of responses “great significance”,
- \(n_D\) – number of responses “big significance”,
- \(n_S\) – number of responses “medium significance”,
- \(n_N\) – number of responses “little significance”,
- \(n_{BN}\) – number of responses “very little significance”,
- \(n_{BZ}\) – number of responses “no significance”;
- \(n\) – total number of responses.
The index ranges from 0 to 100, and a numerical value above 50 indicates a high degree of significance, which increases the closer it is to an index value of 100. Indicators below 50 indicate a low degree of significance of a given entity, and the lower the indicator value, the lower the significance impact [Nazarko, 2013b, p. 106; Ejdys, 2013, p. 112; Dębkowska, 2013, p. 97].

Entities included in the chain network of aesthetic medicine tourism service providers with a significant impact on innovation have been identified based on the structure of networks of institutional links. Their significance assessment and impact on innovation flow was conducted using a scale ranging from 0 to 5, where 0 means the lack of significance and 5 means a direct role in the flow of innovation in aesthetic medicine tourism.

In each of analyzed categories, the foster index ($Wsi$) is determined in terms of influence of service providers on the flow of innovation (developed on the basis of: Nazarko [2013b, p. 106], Ejdys [2013, p. 112], Dębkowska [2013, p. 97]):

$$Wsi = \frac{n}{n_{max}} \times 100\%$$

where:

- $n$ – number of responses,
- $n_{max}$ – maximum number of responses.

The index ranges from 0% to 100%, a numerical value above 50 indicates a high degree of significance, and the closer the index value is to 100 the higher the degree of importance. Indicators below 50 demonstrate a low degree of entity importance, and the closer the indicator value is to zero the lower the significance of its impact [Nazarko, 2013b, p. 106; Ejdys, 2013, p. 112; Dębkowska, 2013, p. 97].

According to the value of the WSI index, direct supporting and auxiliary connections have been distinguished.

Direct links clearly affect the flow of innovations to aesthetic medicine tourism service providers. In immediate relation, the value of the WSI index exceeded 50.00%, although the links supporting this relationship have a lesser impact on the flow of innovations. The value of the WSI index ranges from 49.99% to 25.00% as Auxiliary Relationships only marginally affect the flow of innovations. The WSI index value ranges from 24.99% to 1.00% of indications.

Five categories were used to specify the role of different actors in the innovation process: does not play a role, facilitates the diffusion of innovation, ensures the commercialization of innovation, forms the core of innovation, and participates in the creation of innovation flow.

The results permit the role played by aesthetic medicine tourism in innovation to be identified. Operators directly involved in providing aesthetic medicine tourism services have been indicated, and the institutional links occurring in the innovation flow were defined and their role particularized. The types of ties and the flow directions of innovation were also defined.
Results of Research

There are four main health tourism segments. They are spa tourism, wellness and spa, medical tourism, and aesthetic medicine tourism. According to 41.67% of respondents, study of the latter segment should be expanded.

The provision of aesthetic medicine tourism services is dominated by three groups of service providers: private medical offices (power of influence 66.66), spa treatment establishments (power of influence 51.67) and specialized physiotherapy centers (power of influence 50.00). A lower impact was attributed to the services associated with accommodation (power of influence 40.00) and tourism organizers (power of influence 35.00–33.00). The results are presented graphically in Figure 1.

![Figure 1](image-url)

**Source:** own study based on research results.

Our analysis of institutional links between the service providers in this segment was based on their importance in the flow of innovation, and revealed a complex multi-vector network. We found the strongest linkages between providers of aesthetic medicine tourism services and tour operators. Links between these entities are also directly responsible for the flow of innovation. A detailed analysis of interlinkages indicated the presence of two
internal networks. Each of these networks have shared, separate links that allowed clusters to be identified. Specialized equipment manufacturers, pharmaceutical manufacturers and organizers are the strongest links directly responsible for the flow of innovation in the internal network of aesthetic medicine tourism services providers.

Supporting links within the flow of innovation are ensured by centers for disseminating knowledge, service recipients, providers of accommodations and the human capital of the company. Firms with supporting links are the centers of knowledge creation and suppliers, including the suppliers of specialized equipment and sales agents.

Within the internal network of travel organizers, aesthetic medicine tourism services providers are directly responsible for the flow of innovation. A supporting role in this flow is also played by specialized sale agencies, human capital companies, firms promoting knowledge, recipients of services, and competitors. A complementary role in the flow of innovation is played by accommodation providers and insurance agencies.

The analysis identified types of firms belonging to one network and affecting the flow of innovation in both types of entities. These are providers of accommodation services, centers for the dissemination of knowledge, services recipients, competitors and human capital companies.

FIGURE 2. Institutional relationships in the flow of innovation

Source: own study.
Due to the number of links between institutions in the network, simple single, simple complex, and complex distributed linkages have been identified.

Simple single relationships occurred between two types of providers (aesthetic medicine and accommodation services). Simple complex links occurred through cooperation among several service providers (aesthetic medicine, accommodation, tourist organizing services) whose goals were converging or complementary. In both types of links there is a single direction of flow of innovation.

**FIGURE 3. Flow diagram of innovation**

![Flow diagram of innovation](image)

*Source: own study based on research.*

Complex distributed links characterize the entire network. They result from innovation occurring in various fields of activities. In this network it is not possible to determine the dominant direction of the flow of innovation. Figure 3 presents a flow diagram of innovation resulting from institutional links.

Research concerning the role of firms in creating innovation indicates that private medical offices have the highest significance. Hospitals are at the core of innovation. At the stage of commercialization of innovations, an important role is played by spas, clinics and individuals providing aesthetic medicine tourism services. In the chain of diffusion of innovation, the highest level of importance is attributed to hotels and other accommodation facilities, with less importance given to travel organizers and tourist agencies (Figure 4).
FIGURE 4. **Entities involved in the provision of aesthetic medicine tourism services and their role in the innovation process**

![Diagram showing entities in the innovation process]

- **Facilitates the diffusion of innovation**
- **Ensures the commercialization of innovation**
- **Forms the core of innovation**
- **Participates in the creation of innovation**

Source: own study based on research results.

The above analysis allows for the assignment of roles to firms in the innovation process provided through individual links in the process (Figure 5).

FIGURE 5. **Dominant players in the process of innovation of aesthetic medicine tourism according to the cells of the innovation process**

![Diagram showing dominant players]

Source: own study based on research.
Conclusions and recommendations

The test results lead to the conclusion that aesthetic medicine tourism is an innovative form of health tourism. The most important role in this form of tourism services is played by aesthetic medicine service providers. The test results lead to the conclusion that the purchase of the tourist product, the core of which is aesthetic medicine, is crucial in the aesthetic medicine tourism.

The flow of innovation in this segment occurs in a complex multi-sectoral network. Within it, the largest impact on the innovation originates from aesthetic medicine service providers (private clinics, health treatment facilities, specialized clinics and physiotherapy). Suppliers of medical services also play a most important role in the first three links of the innovation development process (creation, core, and commercialization of innovation). Accommodation and tourist services providers are responsible for the final link in this process – the diffusion of innovation.

One positive phenomenon in the process of developing innovation in aesthetic medicine tourism is the presence of many actors in the providers’ supply chain. The complicated network of institutional entities included in the supply chain can determine the high potential of innovativeness of these services.

The scope of innovation in aesthetic medicine tourism has not yet been tested. We recognize the need for further research to identify the types of innovation that arise in an institutionally complex, multi-vector network of innovation flow. Another potentially interesting research problem involves the innovation system. The practical aspect of the research also requires the identification of indicators of evaluation of innovative activities in this sector.

Notes

1 Author’s e-mail address: e.panfiluk@pb.edu.pl
2 Author’s e-mail address: aleksander.panasiuk@wzieu.pl
3 Author’s e-mail address: e.szymanska@pb.edu.pl
4 The project has been financed through a National Science Centre grant, decision No. DEC-2013/11/B/HS4/02138.
References


Innowacje i transfer technologii (2005), *Słownik pojęć*, PARP, Warsaw.


Rectification

In the previous issue of IJME we mistakenly printed that Tomasz Pruszkowski was affiliated to the Chair of Capital Markets at the Warsaw School of Economics. Correctly the author is a doctoral student at the Collegium of World Economy at the Warsaw School of Economics. We express our apologies to the interested parties.