University of Groningen

Long-term developments of EU household financial product portfolios

Paas, Leonard J.; Bijmolt, Tammo H. A.; Vermunt, Jeroen K.

Published in:
Metron

DOI:
[10.1007/s40300-015-0067-2](https://doi.org/10.1007/s40300-015-0067-2)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Paas, L. J., Bijmolt, T. H. A., & Vermunt, J. K. (2015). Long-term developments of EU household financial product portfolios: a multilevel latent class analysis. *Metron*, 73(2), 249-262.
<https://doi.org/10.1007/s40300-015-0067-2>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Long-term developments of respondent financial product portfolios in the EU: a multilevel latent class analysis

Leonard J. Paas¹ · Tammo H. A. Bijmolt² ·
Jeroen K. Vermunt³

Received: 13 November 2014 / Accepted: 17 June 2015 / Published online: 15 July 2015
© Sapienza Università di Roma 2015

Abstract Segmentation structures can be unstable over time. Therefore, previous research has analyzed panel data in order to provide insight into changes in segmentation structures or switches by individuals between segments. Unfortunately, panel data are often unavailable when analyzing developments across different countries and over longer time-periods. The analysis reported in this paper makes it possible to evaluate differences in segmentation structures across countries at different time-points using multiple cross sectional datasets. This provides indications into long-term cross-national developments of segmentation structures. In the utilized multilevel latent class analysis model respondents are the lower level units and data from the same country and time-point are treated as the higher level units. As an illustrative and salient empirical example we assess similarities and differences in consumer financial product portfolios across 14 EU countries from 1969 to 2003, based on three disaggregate cross-sectional databases.

Keywords Cross-country segmentation · Multilevel latent class analysis · Financial services · Long-term developments · European Union

✉ Leonard J. Paas
L.J.Paas@massey.ac.nz

Tammo H. A. Bijmolt
t.h.a.bijmolt@rug.nl

Jeroen K. Vermunt
j.k.vermunt@uvt.nl

¹ School of Communication, Journalism and Marketing, Massey University,
Private Bag 102904, North Shore Auckland 0745, New Zealand

² Department of Marketing, Faculty of Economics and Business, University of Groningen,
Landleven 5, 9747 AD Groningen, The Netherlands

³ Department of Methodology and Statistics, Tilburg University,
PO Box 90153, 5000 LE Tilburg, The Netherlands

1 Introduction

Segmentation analysis aims to classify units, such as respondents, in categories. However, segmentation structures are often unstable over time, because the classified units may switch between classes and the definitions of classes can change [38]. To accommodate for such dynamics, previous research has applied latent Markov models [3,27] to panel data, in which the same respondent is interviewed at multiple time-points. However, panel data are often unavailable, in particular when analyzing different countries over longer time-periods.

The analysis presented below, derives insights into the similarities and differences between segmentation structures across countries and times, facilitating assessment of long-term cross-country developments. For this, we apply multilevel Latent Class Analysis (LCA) [6,13,34], an extension of the more conventional LCA model [14,19,22]. In conventional LCA there is one type of measurement unit, e.g., individual respondents; multilevel LCA includes: (1) lower level units, e.g., respondents, and (2) higher level units, e.g., countries to which respondents are allocated.

We apply multilevel LCA to analyze disaggregate data on consumer financial product portfolios that occur across 14 EU countries in 1969, 1990 and 2003. At each time-point a different random sample of respondents was interviewed; the individual respondents were the lower-level measurement units. Data from the same country and time-point lead to the definition of higher-level units. For example, Dutch respondents interviewed in 1969 belong to the higher-level unit Netherlands_1969, 1990 respondents from The Netherlands are allocated to Netherlands_1990, etc. This results in 42 higher-level measurement units. A formal definition of the model is provided in Sect. 3.

We analyze consumer financial product portfolios, because in various disciplines empirical studies have been published on this topic, e.g., economic psychology [23,39], statistics [27,28], operations research [29], marketing research [6,18] and services marketing [25]. Insight into consumer financial product portfolios is relevant for marketing activities such as segmentation and cross-selling products [18,26]. In internationalized markets insights into the cross-national similarities and differences between financial product portfolios may reveal fruitful directions for international marketing strategy formulation [6]. To the best of our knowledge only three previous studies have analyzed relevant cross-country data [6,24,28], but these were based on a single cross section. For single country segmentation research, analysis of time dynamics has been accepted as a key element [38], also for consumer financial product portfolios [27,29]. Given the dynamic nature of international financial markets one may anticipate that international segmentation structures, representing respondent financial product portfolios, will change over time. Financial firms require insight into such changes and have to react to international trends [6].

The contribution of the paper is twofold. First and foremost, the paper has a methodological contribution by providing an approach for using multiple international cross-sectional datasets when no longitudinal data are available. This approach can be applied in future empirical studies. Note that the applied model is not new, but the coding of higher-level units based on country in addition to time, is a novel application that potentially provides new insights. Second, the presented multilevel LCA application contributes to knowledge about consumer financial product portfolios and international developments herein.

The organization of the paper, in Sect. 2 we discuss theories that are relevant for cross-national similarities and differences in developments of consumer financial product portfolios. Section 3 presents the applied multilevel LCA model. In the empirical study, reported in Sects. 4 and 5, we assess country specificity of segmentation structures based

on financial product ownership by individual respondents. The paper is concluded with a discussion on theoretical and managerial implications, Sect. 6.

2 International segmentation of financial product portfolios

In the managerially focused marketing literature the analysis of consumers across countries is an important topic [8,32]. Our application of the multilevel LCA model to consumer financial product portfolios is of theoretical interest, as there are different possible outcomes. First of all, we may expect highly similar segmentation structures across different countries. Theory on the development of consumer financial product portfolios within a single country suggests that similar segmentation structures of such product portfolios occur within and across countries. That is, consumer financial product portfolios are considered to be the observable outcomes of the respondent lifecycle and the resulting savings needs that respondents are likely to find relevant [18]. That is, four hierarchically ordered saving motives have been studied extensively in the field of economic psychology [15,20,39]. The most basic level is called the cash management motive, involving short-term financial issues, such as conducting of financial transactions. At level two, represented by the precautionary motive, respondents develop a financial reserve for unexpected expenditures. This is followed by the down-payment motive at level three, i.e., accumulating a financial deposit for a house, saving for a car, an extensive vacation or durable products. At the fourth level, wealth management consists of enterprise and investing assets. The second construct, the lifecycle hypothesis and the related permanent income hypothesis [7], assumes acquisitions result from consumer circumstances, such as the lifecycle phase and income. Young consumers require financial products for borrowing or investing small amounts of assets. Later in life, when income and assets increase, consumers require more sophisticated products for reasons such as speculation and asset accumulation.

Kamakura et al. [18] suggest that the lifecycle theory and saving needs are interrelated and lead to a common order for acquiring financial products. Such an acquisition order reflects: (1) assets that are invested and (2) risk levels of the different products. Consumers in early life-cycle stages have higher priorities for financial products related to basic motives in the saving motive hierarchy, involving fewer assets and relatively low risk levels, e.g., saving accounts. Products that are relevant for higher order motives, e.g., shares, will often be acquired by consumers with more assets and higher levels of financial knowledge. We conjecture that the tendency of consumers to acquire more basic financial products ahead of products satisfying higher order needs may occur across countries. The latter would imply that consumers in different countries tend to own higher order products when they also own products relevant for the satisfaction of more basic needs, resulting in similar financial product-portfolios across countries, i.e., similar segmentation structures.

Contradicting expectations are also feasible, due to country-specific situational factors. Cultural factors may be particularly relevant herein, such as those captured through the well-known Hofstede dimensions [17]. For example, cross-country differences on the uncertainty avoidance dimension in Hofstede's model may be reflected in risk aversion of consumers in the financial market. In low uncertainty avoidance cultures the consumer may less often develop a financial buffer in safe assets, such as in savings accounts, but may first choose to invest in risky assets, such as shares. Cultural differences, involving uncertainty avoidance and also the other Hofstede dimensions [11], persist between EU countries, the region that is analyzed in our empirical study. Another relevant issue concerns GDP differences across countries. Extant research has consistently shown that consumers with higher income levels

are more likely to own a larger variety of financial products [7,39]. We conjecture that this may also be reflected in the differences between consumer financial product portfolios across countries with varying GDP's.

In sum, the lifecycle hypothesis [7] and the savings need hierarchy [15,20,39] would suggest that the segmentation structures derived from consumer financial product portfolios are likely to be similar across countries. Contrary to this, the literature on persisting cultural differences across countries in the EU [17] and GDP differences would suggest occurrence of divergent segmentation structures across EU countries. This contradiction implies that the application of the multilevel LCA presented below is theoretically relevant.

3 Model specification

In this paper a multilevel LCA model is employed in a novel way. Bijmolt et al. [6] and Paccagnella and Varriale [28] previously used this model for assessing financial product portfolios across countries at one time-point. We apply multilevel LCA to data collected in multiple countries at various time-points, as discussed in Sect. 1.

We define the model using the notation in Bijmolt et al. [6]. Assume that data are available on multiple respondents, denoted $i = 1, \dots, I$. The I respondents are from a set of countries, denoted $j = 1, \dots, J$, and were interviewed at various time-points, $t = 1, \dots, T$. For each respondent i , product ownership indications are available on a set of products, denoted $k = 1, \dots, K$, where $Y_{ijtk} = 1$, if respondent i from country j interviewed at time-point t owns product k , else $Y_{ijtk} = 0$. All ownership indications for respondent i are represented in the vector Y_{ijt} . Furthermore, Y_{jt} is a matrix containing ownership data of all respondents of country j at time-point t . The model assumes S respondent-level latent classes, denoted $s = 1, \dots, S$. The analyzed countries belong to higher level latent classes at the different time-points, denoted $h = 1, \dots, H$. In our application, the dataset analyzed contains information from 14 countries at three time-points resulting in 42 time-by-country combinations. These 42 higher level units are placed in H higher-level latent classes. The categorical latent variable X_{ijt} represents respondent latent class memberships and Z_{jt} represents higher level latent class memberships. Note that we formulate a two-level model involving all possible time-country combinations. It would also be possible to formulate a three-level model, with times denoting the second level and countries the third level. A two-level formulation facilitates a more flexible allocation of countries at different points in time to the same higher-level segment. Furthermore, the number of higher level units in our database is too limited for a three-level model.

Multilevel LCA models [6,34] consist of two model equations, one at the respondent level (Eq. 1) and another at the higher level, i.e., the country-by-time level in our study (Eq. 2). For respondents, probabilities of product ownership, Y_{ijtk} , for respondent i from country j at time-point t , is conditional on membership of country j at time-point t to the higher-level latent classes, h , i.e.:

$$P(Y_{ijtk} | Z_{jt} = h) = \sum_{s=1}^S P(X_{ijt} = s | Z_{jt} = h) \prod_{k=1}^K P(Y_{ijtk} | X_{ijt} = s). \quad (1)$$

Equation (1) can be interpreted as a regular LCA model. However, in this multilevel LCA equation the relative sizes of the S latent classes at the respondent level are affected by the higher-level latent class memberships of the 42 higher-level units. As in the conventional LCA model it is assumed that within a latent class s , the ownership probabilities of the various financial products are independent, i.e., local stochastic independence.

At the country-by-time level, a second equation is specified:

$$P(Y_{jt}) = \sum_{h=1}^H P(Z_{jt} = h) \prod_{i=1}^{N_{jt}} P(Y_{ijt} | Z_{jt} = h), \quad (2)$$

where N_{jt} represents the sample size of country j as interviewed at time-point t . Combining Eqs. (1) and (2) leads to the following:

$$(P(Y_{jt})) = \sum_{h=1}^H \left[P(Z_{jt} = h) \prod_{i=1}^{N_{jt}} \left[\sum_{s=1}^S P(X_{ijt} = s | Z_{jt} = h) \prod_{k=1}^K P(Y_{ijk} | X_{ijt} = s) \right] \right] \quad (3)$$

In Eq. (3) the outcome $P(Y_{jt})$ results from: (a) the probability that country j at time-point t belongs to the higher-level latent class h , (b) the probability respondent i belongs to consumer latent class s , given the higher-level latent class membership, h , and (c) the probability respondent i owns product k , given the respondent latent class membership, s . Component (c) of Eq. (3) involves the differences between the S respondent-level latent classes, which are the conditional probabilities that a respondent owns a product k , given this unit's segment membership. For parameter estimation, these probabilities are parameterized through a logit equation:

$$P(Y_{ijk} = 1 | X_{ijt} = s) = \frac{\exp(\beta_{ks})}{1 + \exp(\beta_{ks})}, \quad (4)$$

where β_{ks} is the logit of ownership of product k by individuals belonging to class s . Component (b) of Eq. (3) assesses key differences between the H higher-level latent classes. This is based on the relative sizes of respondent-level latent classes in country j at time point t , i.e.:

$$P(X_{ijt} = s | Z_{jt} = h) = \frac{\exp(\gamma_{sh})}{\sum_{s'=1}^S \exp(\gamma_{s'h})}, \quad (5)$$

where γ_{sh} is a multinomial logit coefficient. We also anticipate effects of demographics on respondent latent class membership probabilities. These covariate effects are modeled using the concomitant variables approach [1, 9, 12], which was applied previously by Bijmolt et al. [6] for retrieving covariate effects in multilevel LCA. The model for the lower-level classes becomes:

$$P(X_{ijt} = s | Z_{jt} = h, W_{ijt}) = \frac{\exp(\gamma_{sh0} + \sum_{p=1}^P \gamma_{sp} W_{ijtp})}{\sum_{s'=1}^S \exp(\gamma_{s'h0} + \sum_{p=1}^P \gamma_{s'p} W_{ijtp})}, \quad (6)$$

where W_{ijtp} is one the P covariates and γ are logistic regression parameters.

Model parameters can be estimated through maximum likelihood using an Expectation-Maximization (EM) algorithm [5, 10]. For this purpose, a special implementation of the E step of the EM algorithm is required, which was described in detail by Vermunt [34–36], who called it an upward-downward algorithm. In the upward step, one obtains the posterior probabilities for the higher-level class memberships based on all information of the higher-level country-time unit concerned; that is, $P(Z_{jt} = h | Y_{jt})$. In the downward step, one obtains the bivariate posteriors $P(Z_{jt} = h, X_{ijt} = s | Y_{jt})$ using $P(Z_{jt} = h | Y_{jt})$ and $P(X_{ijt} = s | Z_{jt} = h, Y_{ijt})$. Using these posterior probabilities one can construct the expected complete data log-likelihood which is maximized in the M step of the EM algorithm. For the latter, standard Newton-type algorithms for logistic regression can be used. We used

the Latent GOLD 5.0 program [37] for our analysis. Note that the R package MultiCIRT [4] also implements the multilevel LCA model.

Reweighting was required because the analyzed cross-country respondent-level data are based on national samples that are disproportional for the actual population size. That is, in the sample a larger percentage of respondents were selected from those countries with smaller populations. For example, the 2003 dataset includes 493 respondents from Luxembourg with a population of approximately half a million. The number of respondents from Great Britain is approximately double that of Luxembourg in 2003, namely 1054. This is not proportionate to the larger population of Great Britain, almost 60 million in 2003. In a previously reported application of multilevel LCA, Bijmolt et al. [6] employed the reweighting procedure suggested by Ter Hofstede et al. [33], in which respondents from countries with smaller populations are down-weighted, while respondents from countries with larger populations are up-weighted. In the current paper, we use a different kind of weighting which fits better to the multilevel nature of our data and model. It involves using weights at both the respondent level and the higher country-time level. The approach reweights (down-weights) the individual-level data in such a manner that each country-time combination contains the equivalent of 100 respondents, which corresponds with a multilevel analysis with rather large (and equally sized) groups. At the country level, weights are applied to accommodate for population size differences between countries and time points: Countries with larger populations receive higher weights, where for each country-time combination the sum of the product of the lower- and higher-level weights are proportional to the population sizes. Similar weighting schemes with products of multilevel weights proportional to the total number of lower-level units per higher-level units in the population have been proposed for other types of multilevel level models by Asparouhov and Muthén [2] and Rabe-Hesketh and Skrondal [30].

Relative fit of alternative model specifications is evaluated using the minimum Bayesian Information Criterion (BIC) rule [31]. In the comparison of the relative fit of alternative models we use the three-step procedure introduced by Lukociene et al. [21]. In the first step, one estimates a conventional LCA model [14, 19] with a single level of latent classes for segmenting respondents. Here we use the BIC based on the number of individuals for selecting an optimal number of segments. Next the number of lower-level latent classes is fixed to the value that has been attained through step 1 and the appropriate number of higher-level classes is determined using the minimum BIC rule, but now using the number of country-time combinations as the sample size (step 2). In the final step 3 the number of higher-level classes is fixed to the value attained in step 2 and the number of lower-level classes is determined again using BIC based on the lower-level sample size. To account for sub-optimal solutions we estimated each model 160 times with different random starting values, retaining the best solution.

4 Data and analysis

The 1969 and 1990 data were collected for the EURODATA project of the Readers Digest Respondent Survey. The 1969 sample contains 24,180 respondents from 16 countries and the 1990 sample 22,339 individuals from 17 countries. Respondents were 18 years or older and were living in private households. The 2003 data, Eurobarometer 60.2, were collected by a consortium of market research agencies at request of the European Commission, Directorate-General Press and Communication, Opinion Polls. It covers the population (aged 15 years and over) of the 15 EU member states in 2003 and consists of 16,200 respondents. In all years, interviews were conducted face-to-face and in the appropriate national language.

Table 1 Product penetrations over time

Product	1969	1990	2003
<i>Investments</i>			
(1) Shares	0.05	0.09	0.17
(2) Investment trusts	0.04	0.03	0.15
(3) Bonds	0.12	0.06	0.08
<i>Long-term contractual saving</i>			
(4) Life insurance	0.25	0.21	0.43
<i>Regular saving</i>			
(5) Savings account	0.40	0.81	0.90

The empirical study includes those 14 countries that are represented at all three time-points. Respondents aged between 15 and 17 were excluded from the 2003 dataset. This leads to a database with 51,674 respondents, aged 18 years and older, from Austria, Belgium, Denmark, Finland, France, West Germany, Great Britain (including Northern Ireland), Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain and Sweden. Thus, the analyzed dataset covers all countries that joined the EU prior to the ten-country extension in 2004, except Greece and former East Germany. We reweighted the individual observations to equal 100 in each country-time based higher level unit, as outlined in Sect. 3.

Because we analyzed financial product portfolios in 1969, 1990 and 2003, we can assess country-level developments in respondent financial product portfolios. Note that a long time-span is analyzed, because effects of internationalization are expected to take place over longer periods [11]. Also, financial product portfolios are developed over long-term respondent lifecycles [7,25,39]. The disaggregate nature of the dataset implies insight is provided into similarities and differences across respondents within the same country and also between respondents across countries [6] and across measurement occasions.

We analyzed the following products: savings account, life insurance, bonds, investment trusts, and shares. Table 1 presents the time-specific penetrations of these financial products, ranked according to decreasing risk levels [23]. Shares are the most risky assets, depending on highly volatile stock market exchange values. Investment trusts involve a mix of different shares and corporate/government bonds and are subjected to smaller value fluctuations. Next are the corporate/government bonds. Life insurances are less risky again. Like investment trusts these concern a mix of shares and bonds. However, investments are usually over longer periods and, therefore, involve lower risks [23]. Least risky are the savings account.

The database also includes three respondent-level demographics: age, income and marital status. Previous research has shown these variables to be of key importance for consumer financial product portfolios [7,21,39]. The demographic variables are used to profile the financial product portfolio based segments derived from the dataset, by being included in the model as covariates.

5 Empirical results

5.1 Model selection

As discussed above, we employed the three-step procedure [21] for determining an appropriate number of lower- and higher-level latent classes. Table 2 presents the BIC values for

Table 2 Model selection in the three steps of the Lukociene et al (2010) procedure: Log-likelihood values (LL), number of parameters (Npar), and BIC values

Step	Higher level (H)	Lower level (S)	LL	Npar	BIC, N = 4200	BIC, N = 42
1	1	1	-8773	5	17,588	
1	1	2	-7858	16	15,849	
1	1	3	-7733	27	15,692	
1	1	4	-7678	38	15,672	
1	1	5	-7629	49	15,666	
1	1	6	-7602	60	15,704	
2	1	5	-7629	49		15,440
2	2	5	-7084	54		14,369
2	3	5	-6980	59		14,181
2	4	5	-6926	64		14,092
2	5	5	-6881	69		14,021
2	6	5	-6851	74		13,979
2	7	5	-6831	79		13,957
2	8	5	-6817	84		13,948
2	9	5	-6807	89		13,946
2	10	5	-6798	94		13,948
3	9	4	-6911	70	14,406	
3	9	5	-6807	89	14,356	
3	9	6	-6777	108	14,455	

the alternative models in each step. In the first step, which ignores the multilevel structure in the LCA model, we find that the LCA with five latent classes results in the lowest BIC. Consequently, in the second step, the number of lower-level latent classes is fixed to five and the number of higher-level latent classes is increased from 1 to 10, while estimating the multilevel LCA. The model with nine higher-level classes leads to the lowest BIC, see Table 2. In step 3 the number of higher-level latent classes is set to nine and the number of lower-level latent classes is again investigated using multilevel LCA. For this purpose, we estimated models with one class less and one class more than five. We find again that five lower-level latent classes result in the lowest value on BIC. Therefore, the model with five lower level latent classes and nine higher level classes has been selected.

5.2 The consumer segmentation solution

Table 3 presents latent class specific product penetrations for the lower level segmentation structure involving respondents. In the first latent class, ‘in-actives’ ($s = 1$), penetrations of all products are close to zero. In the second class, ‘savers’ ($s = 2$), only the savings account has a high penetration. In the third class, ‘life-savers’ ($s = 3$), respondents own the savings account and the life insurance, but not the three other financial products. In the fourth class, ‘life-savers with bonds’ ($s = 4$), we find respondents that generally own a savings account, and relatively often have life insurance and bonds. In the fifth latent class, i.e., ‘actives’ ($s = 5$), are respondents that often own savings accounts, life insurances, investment trusts and shares. Sometimes they also own bonds.

Table 3 Segment-specific product penetrations, $P(Y_{ijtk}|X_{ijt} = s)$

Segment label	In-actives $s = 1$	Savers $s = 2$	Life-savers $s = 3$	Life-savers with bonds $s = 4$	Actives $s = 5$
Cluster size	0.2526	0.4104	0.1669	0.0726	0.0974
Shares	0.0014	0.0315	0.0989	0.1337	0.6547
Invest. trusts	0.0004	0.0127	0.0463	0.0337	0.6530
Bonds	0.0017	0.0150	0.0107	0.5042	0.3703
Life insurance	0.0278	0.0387	0.9845	0.6718	0.6584
Savings account	0.0144	0.9642	0.9629	0.9196	0.9860

Table 4 Covariate effects on respondent level segment membership

Segment label	In-actives ($s = 1$)	Savers ($s = 2$)	Life-savers ($s = 3$)	Life-savers with bonds ($s = 4$)	Actives ($s = 5$)	Sign.
Partner	-0.18	-0.02	0.26	-0.23	0.16	df = 4, Wald = 22.02, $p < 0.01$
Income > median	-0.24	-0.21	-0.02	-0.14	0.61	df = 4, Wald = 68.13, $p < 0.01$
Age < 35	0.53	0.31	0.05	-0.28	-0.62	df = 12
Age 36–50	-0.20	-0.05	0.39	-0.17	0.04	Wald = 102.55
Age 51–64	-0.29	-0.16	0.06	0.03	0.36	$p < 0.01$
Age 65+	-0.04	-0.10	-0.49	0.42	0.22	

The results reported in Table 3 suggest that respondents have followed similar orders in acquiring financial products. That is, in each latent class where the other four products have high penetrations the savings account is also commonly owned, implying that respondents usually acquire savings accounts before these other products [6, 27]. They are likely to acquire life insurance at a later stage, because this product has high penetrations in classes where the bonds, investment trusts and shares also are commonly owned. In $s = 4$ we find a high penetration for bonds and lower penetrations for investment trusts and shares. By contrast in latent class $s = 5$ we find higher penetrations for investment trusts and shares than for bonds. This difference implies some respondents owning a savings account and life insurances chose between first acquiring bonds, on the one hand, or investments trust and/or shares, on the other hand.

Table 4 reports the covariate effects of the demographics, i.e., partner (y/n), age category and income above median (y/n), on the latent class membership probabilities of respondents. The effects of income, age and having a partner are all significant ($p < 0.01$). Respondents with an income above the median are less often in the segments in which few products have high penetrations, i.e., in-actives ($s = 1$) and savers ($s = 2$). This also applies for respondents with a household head of intermediate or higher age (51+) and for respondents in which the household head has a partner. These findings are mostly consistent with extant theory and previous research [6, 35] and thereby support validity of our results.

5.3 Clustering of Countries over time

Table 5 presents the H higher-level latent classes which represent the commonly occurring combinations of the S lower-level latent classes within the 42 different higher-level units. We find that some higher-level latent classes are characterized by respondent level latent class structures that are not very advanced in terms of financial product portfolios. For example, higher-level latent class $h = 1$ contains countries at a time that many of the respondents (79.33 % on average) are in the in-active segment $s = 1$. Contrarily, if at a specific time-point t a country has many respondents that are actives ($s = 5$) the corresponding segmentation structure of the country at time t is more likely to be allocated to, for example $h = 9$. To enhance interpretation of Table 5 we have ordered the nine higher-level latent classes according to decreasing proportions of the inactive respondent-level, i.e., respondents-level latent class $s = 1$ and increasing proportions of the active respondents, i.e., $s = 5$.

As mentioned above, allocation of the 42 higher-level units in higher-level latent classes is based on the occurrence of the five respondent-level latent classes. The respondent-level latent classes represent the lower level segmentation structure in the multilevel LCA model. For example, respondent level latent class $s = 1$ in Table 3 is defined by an average ownership probability of 0.0014 for shares, 0.0004 for investment trusts, 0.0017 for bonds, 0.0278 for the life insurance product and 0.0144 for the savings account. If the occurrence of the five respondent level latent classes is highly similar across different higher-level units, these higher-level units are likely to be allocated to the same country-by-time latent class, h . An illustrative hypothetical example to consider is, that in 1969, we find 25 % of the Belgian respondents in lower-level latent class $s = 1$, 50 % of the respondents in this country is in $s = 3$ and the remaining 25 % is in $s = 4$. Moreover, in France (1969) we find a highly similar segmentation structure, with 20 % of the respondents in $s = 1$, 60 % in $s = 3$ and the remaining 20 % is in $s = 4$. Under such circumstances Belgium (1969) and France (1969) are likely to be allocated to the same higher level latent class h . However, if two or more higher-level units are characterized by dissimilar segmentation structures, they would probably be allocated to different higher level classes. Building on our hypothetical example, consider that in Great Britain in 2003 we find 50 % of the respondents in $s = 2$ and the other 50 % in $s = 5$. In this scenario Great Britain (2003) is unlikely to be allocated to the same higher level latent class h as Belgium (1969) and France (1969). Segmentation structures of countries at different time-points can also be allocated to the same higher level class. For example, the respondent level segmentation structure of Portugal in 2003 is highly similar to

Table 5 Definition of the higher-level latent classes, $P(X_{ijt} = s | Z_{it} = h)$

Higher-level class number	$s = 1$	$s = 2$	$s = 3$	$s = 4$	$s = 5$
$h = 1$	0.7933	0.1368	0.0437	0.0103	0.0159
$h = 2$	0.5965	0.1881	0.1193	0.0744	0.0217
$h = 3$	0.5630	0.0007	0.0134	0.4038	0.0190
$h = 4$	0.3265	0.5815	0.0310	0.0394	0.0218
$h = 5$	0.0884	0.6290	0.1849	0.0517	0.0459
$h = 6$	0.1824	0.4507	0.0947	0.0598	0.2124
$h = 7$	0.0239	0.5036	0.3040	0.0040	0.1644
$h = 8$	0.0792	0.3314	0.3417	0.0763	0.1714
$h = 9$	0.0261	0.1216	0.1463	0.0059	0.7001

Table 6 Time-specific allocation of countries to the higher-level segments

Country	1969	1990	2003
Austria, Belgium, Luxembourg, Finland, France, The Netherlands	2	5	7
Portugal, Spain	1	4	5
Italy	1	4	6
Ireland	3	4	8
GB, Denmark	3	5	8
Sweden	3	5	9
West Germany	3	5	7

the structure characterizing Denmark in 1990. Therefore, Portugal-2003 and Denmark-1990 are placed in the same higher-level latent class $h = 5$, see Table 6. This suggests that Portugal is lagging in terms of the development of consumer financial product portfolios and Denmark is more advanced.

Table 6 reports the complete allocation of the 14 countries in 1969, 1990 and 2003 to the nine higher-level latent classes. Each of the 42 higher-level units was allocated to one of the H higher level latent classes using the modal classification rule; that is, to the class for which the posterior membership probability $P(Z_{jt} = h|Y_{jt})$ is largest. An obvious observation that can be derived from Table 6 is that financial product portfolios have become more divergent across the 14 analyzed countries. In 1969 and 1990 these countries were allocated to only two and three of the higher-level latent classes. This increased to five higher-level latent classes in 2003.

A second finding is that some sets of countries display similar developments. This is reflected in the same or highly similar allocation of countries to the higher-level latent classes across the three analyzed time-points. Foremost, for six of the analyzed countries this allocation is the same: (1) Austria, Belgium, Luxembourg, Finland, France and The Netherlands. These countries were in higher-level latent class $h = 2$ in 1969, $h = 5$ in 1990 and $h = 7$ in 2003.

Next to this Table 6 shows that the three Mediterranean countries, Italy, Portugal and Spain, also underwent similar developments, being allocated to higher-level latent class $h = 1$ in 1969. They proceeded to $h = 4$ in 1990 and then Portugal and Spain entered $h = 5$ in 2003, which also includes some of the more advanced North-European markets in 1990. The lagging of two of the three Mediterranean countries suggests using successful marketing strategies and tactics previously applied in the more advanced markets. Italy was allocated to $h = 6$ in 2003, a higher-level latent class that doesn't include any other countries at any time-point, implying that in 2003 Italy had a relatively unique segmentation structure.

A third set of countries, GB, Ireland and Denmark, started in $h = 7$ in 1969 and proceeded to $h = 8$ in 2003, via different higher-level latent classes in 1990. That is, GB and Denmark were in $h = 5$ in 1990 and Ireland was in $h = 4$, with the Mediterranean countries. The economic boom of the 1990's may explain why Ireland has developed rapidly between 1990 and 2003 in terms of consumer financial product portfolios, catching up with more developed North European countries.

West-Germany and Sweden are two countries displaying divergent developments. Perhaps conditions specific to these countries have resulted in their unique developmental pattern. For example, Sweden is allocated to the higher-level cluster $h = 9$ which represents an advanced segmentation structure, see Table 3. No other countries are allocated to $h = 9$ at any measurement occasion. This finding may be due to the Swedish pension scheme in which

all citizens in paid employment are provided with the opportunity to allocate part of their savings to equity funds and other risk-bearing securities [16].

In sum, although some countries display similar patterns in development, countries do not all follow the same developmental process in terms of the financial product portfolios that consumers do. Segmentation structures develop in different ways and at different speeds. These differences have resulted in an increased diversity in segmentation structures across the 14 analyzed countries in 2003.

6 Discussion

In this paper we have illustrated the application of the multilevel LCA model for deriving insights into cross-national longitudinal developments by analyzing multiple cross sectional datasets. In the absence of cross-national longitudinal data our approach can provide relevant insight into such developments at the country level. The application of the multilevel LCA model reported in the current paper has led to some interesting theoretical findings. These will be discussed first then we will conclude with a discussion on methodological implications.

We find that consumers in the 14 analyzed EU countries generally acquire products according to increasing risk levels, i.e., less risky products before more risky products. The suggested pattern does not seem to be disturbed by contextual differences between countries. Nevertheless, we find that countries are developing along this common pattern at different rates. For example, the Mediterranean countries Italy, Portugal and Spain, are lagging behind the other countries that are represented in the analysed dataset. Such divergences support the relevance of persisting cultural and GDP differences across the analyzed EU countries, as discussed in Sect. 2. These empirical results also have managerial implications for financial firms in the EU. First, the similar developmental order of consumer financial product portfolios across countries implies that in all countries respondents owning only a savings account can be offered a life insurance product first, while those respondents that already own a life insurance product can be offered an investment product. In other words, similar marketing strategies and tactics can be developed in such countries. Moreover, marketers can consider using market strategies in lagging countries, e.g., Spain and Portugal, which have proven to be successful in the more advanced countries. Our empirical findings reveal that there is need for further research on other related services and products. Avenues for future research involve two important limitations of our study. First, that we studied just 14 EU countries and found substantial differences amongst these similar countries. Investigating a larger variation of countries would be interesting and beneficial, particularly if developing countries or the new EU members (joining in 2004 or later) and Greece was included. Second, further research could aim to investigate the effects of country-level variables on developments in respondent financial product portfolios.

From a methodological perspective, we have illustrated the application of the multilevel LCA model for analyzing multiple cross-sectional datasets to assess differences between countries and across time-points, which can be used to assess developments in countries and differences in developments across different countries. In the commonly occurring absence of truly longitudinal cross-national datasets this is a relevant application of the multilevel LCA model that can also be used in future research on respondent financial product portfolio development and also for other substantive applications such as developments in ownership combinations of consumer durable products, developments in consumer's personal values across nations or social-demographics developments across nations.

References

1. Andrews, R.L., Currim, I.S.: Recovering and profiling the true segmentation structure in markets: an empirical investigation. *Int. J. Res. Mark.* **20**, 177–192 (2003)
2. Asparouhov, T., Muthén, B.: Multilevel modeling of complex survey data. In: *Proceedings of the Joint Statistical Meeting in Seattle, August 2006. ASA section on Survey Research Methods*, pp. 2718–2726 (2006)
3. Bartolucci, F., Farcomeni, A., Pennoni, F.: *Latent Markov Models for Longitudinal Data*. Chapman and Hall / CRC Press, Boca Raton (2013)
4. Bartolucci, F., Bacci, S., Gnaldi, M.: MultiCIRT: An R package for multidimensional latent class item response models. *Comput. Stat. Data Anal.* **71**, 971–985 (2014)
5. Baum, L.E., Petrie, T., Soules, G., Weiss, N.: A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Stat.* **41**, 164–171 (1970)
6. Bijmolt, T.H.A., Paas, L.J., Vermunt, J.K.: Country and respondent segmentation: Multilevel latent class analysis of financial product ownership. *Int. J. Res. Mark.* **21**, 323–340 (2004)
7. Browning, M., Lusardi, A.: Respondent Savings: Micro Theories and Micro Facts. *J. Econ. Lit.* **34**, 1797–1855 (1996)
8. Craig, C.S., Douglas, S.P.: *Cross-country Marketing Research*, 3rd edn. Wiley, New York (2005)
9. Dayton, C.M., Macready, G.B.: Concomitant-variable latent-class models. *J. Am. Stat. Assoc.* **83**, 173–178 (1988)
10. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. R. Stat. Soc. Ser. B* **39**, 1–38 (1977)
11. De Mooij, M., Hofstede, G.: Convergence and divergence in respondent behavior: Implications for international retailing. *J. Retail.* **78**, 61–69 (2002)
12. Formann, A.K.: Mixture analysis of multivariate categorical data with covariates and missing entries. *Comput. Stat. Data Anal.* **51**, 5236–5246 (2007)
13. Gnaldi, M., Bacci, S., Bartolucci, F.: A multilevel finite mixture item response model to cluster examinees and schools. *Adv. Data Anal. Classif.* (2015). doi:[10.1007/s11634-014-0196-0](https://doi.org/10.1007/s11634-014-0196-0)
14. Goodman, L.A.: Exploratory latent structure analysis using identifiable and unidentifiable models. *Biometrika* **61**, 215–231 (1974)
15. Gunnarson, J., Wahlund, R.: Respondent financial strategies in Sweden: an exploratory study. *J. Econ. Psychol.* **18**, 211–233 (1997)
16. Hedesström, T.M., Svedsäter, H., Gärling, T.: Identifying heuristic choice rules in the Swedish premium pension scheme. *J. Behav. Financ.* **5**, 32–42 (2004)
17. Hofstede, G.: *Cultures consequences: Comparing values, behaviors, institutions and organizations across nations*, 2nd edn. Sage Publications, Thousand Oaks (2001)
18. Kamakura, W.A., Ramaswami, S.N., Srivastava, R.K.: Applying latent trait analysis in the evaluation of prospect for cross-selling financial products. *Int. J. Res. Mark.* **8**, 329–349 (1991)
19. Lazarfeld, P.F., Henry, N.W.: *Latent Structure Analysis*. Houghton Mifflin, Boston (1968)
20. Lindqvist, A.: A note on the determinants of respondent saving behaviour. *J. Econ. Psychol.* **1**, 39–57 (1981)
21. Lukociene, O., Varriale, R., Vermunt, J.K.: The simultaneous decision(s) about the number of lower- and higher-level classes in multilevel latent class analysis. *Sociol. Methodol.* **40**, 247–283 (2010)
22. Magidson, J., Vermunt, J.K.: Latent class models. In: Kaplan, D. (ed.) *The sage handbook of quantitative methodology for the social sciences*, pp. 175–198. Sage Publications, Thousand Oaks (2004)
23. Ngwenya, M., Paas, L.J.: Lifecycle effects on consumer financial product portfolios in South Africa: an analysis of four ethnic groups. *J. Econ. Psychol.* **33**, 8–18 (2012)
24. Paas, L.J.: Acquisition patterns of products facilitating financial transactions: a cross-country investigation. *Int. J. Bank Mark.* **19**, 266–275 (2001)
25. Paas, L.J.: Acquisition pattern analysis for evolutionary database marketing. *Serv. Ind. J.* **29**, 805–812 (2009)
26. Paas, L.J.: Comment on: Latent Markov models: a review of a general framework for the analysis of longitudinal data with covariates. *Test* **23**, 473–477 (2014)
27. Paas, L.J., Vermunt, J.K., Bijmolt, T.H.A.: Discrete time, discrete state latent Markov modelling for assessing and predicting respondent acquisitions of financial products. *J. R. Stat. Soc. A* **170**, 955–974 (2007)
28. Paccagnella, O., Varriale, R.: Asset ownership of the elderly across Europe: A multilevel latent class analysis to segment country and respondent. *Adv. Theor. Appl. Stat.* 383–393 (2013)
29. Prinzie, A., van den Poel, D.: Investigating purchasing patterns for financial services using Markov, MTD and MTDg models. *Eur. J. Oper. Res.* **170**, 710–734 (2006)

30. Rabe-Hesketh, S., Skrondal, A.: Multilevel modeling of complex survey data. *J. R. Stat. Soc. Ser. A* **169**(4), 805–827 (2006)
31. Schwarz, G.: Estimating the dimensions of a model. *Ann. Stat.* **6**, 461–464 (1978)
32. Steenkamp, J.-B.E.M.: Moving out of the US silo: a call to arms for conducting cross-country marketing research. *J. Mark.* **69**, 6–8 (2005)
33. Ter Hofstede, F., Steenkamp, J.-B.E.M., Wedel, M.: Cross-country market segmentation based on respondent-product relations. *J. Mark. Res.* **36**, 1–17 (1999)
34. Vermunt, J.K.: Multilevel latent class models. *Sociol. Methodol.* **33**, 213–239 (2003)
35. Vermunt, J.K.: A hierarchical mixture model for clustering three-way datasets. *Comput. Stat. Data Anal.* **51**, 5368–5376 (2007)
36. Vermunt, J.K.: Latent class and finite mixture models for multilevel datasets. *Stat. Methods Med. Res.* **17**, 33–51 (2008)
37. Vermunt, J.K., Magidson, J.: Technical Guide for Latent GOLD 5.0: Basic, Advanced, and Syntax. Statistical Innovations Inc, Belmont (2013)
38. Wedel, M., Kamakura, W.A.: Market Segmentation: Conceptual and Methodological Foundations, 2nd edn. Kluwer, Dordrecht (2000)
39. Wörneryd, K.-E.: The Psychology of Saving: A Study of Economic Psychology. Edward Elgar Publishing, Northampton (1999)