



ACTUARIAL SOCIETY
OF SOUTH AFRICA

Advancing Actuarial Excellence

Predicting Retirement Income Needs

A Macro-Level Analysis Using Bank Transactional Data

Dr David Rodda, Phuti Malepeng, Nkosinathi Sikhonde

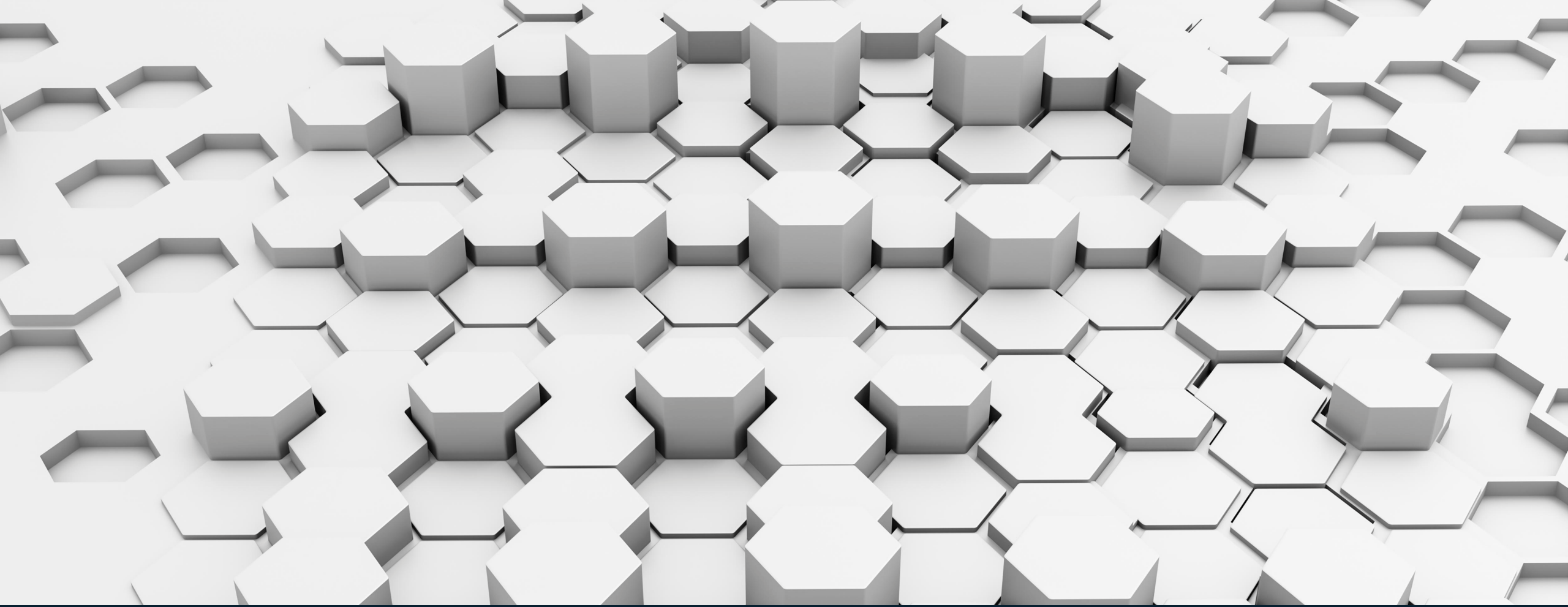
28 July 2025 Retirement Matters Seminar



Agenda

1. Introduction to the Problem and Data
2. Age-period-Cohort Effects
3. Results





Introduction to the Problem and Data






ACTUARIAL SOCIETY
OF SOUTH AFRICA

Advancing Actuarial Excellence

Educate • Develop • Uphold



Introduction to the problem

- Savings targets are back solved from required monthly income targets
- Widespread “rules of thumb” suggest a 70-80% gross replacement rate (MacDonald & Moore 2011) 
- Some research suggests that non-healthcare expenses are more likely to remain unchanged into retirement and healthcare costs may increase (Butler & van Zyl, 2012) 
- Other research indicates that consumption decreases steadily during retirement (Hurd & Rohwedder, 2023) 
- The contrast illustrates that it is difficult to know how much one should expect to need in retirement, making it hard to best prepare.

Lack of consensus on retirement spend trajectories

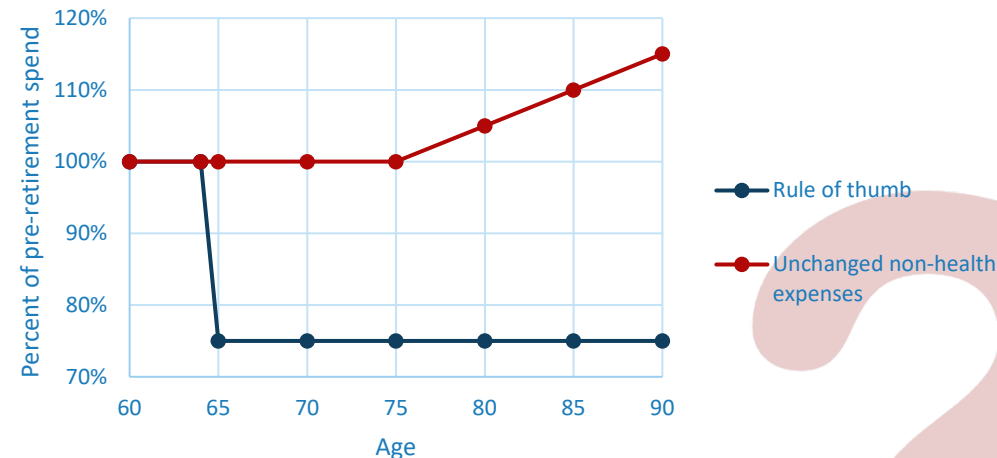
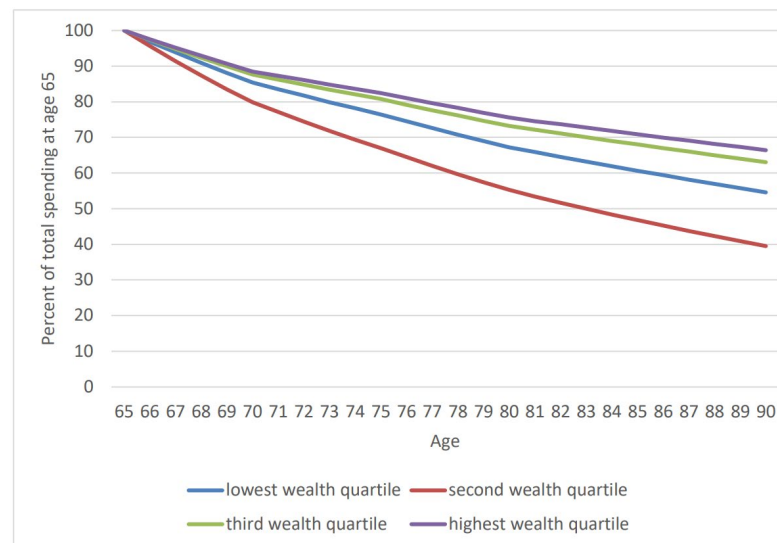


Figure S.1. Estimated Spending Trajectories for Single Households





The Bank Transaction Data Analysed

- National surveys have limited sample sizes, particularly at higher income levels; and rely on self-reporting which can be inaccurate.
 - Discovery Bank provided aggregated data on over 300,000 clients, almost all in the top two income deciles
 - By comparison, the LCS 2014/2015 survey has data for some 22,000 households, and less than 100,000 individuals, of which few are in the top decile of income earners.
-



Pros and Cons of using Discovery Bank Transactional Data

Pros

- Larger sample of individuals (esp. high income)
- Doesn't rely on memory or limited periods diaries
- Abundant extra information
 - Vitality status
 - family composition
 - property and savings values
 - verified income values
 - product
- Characteristics relatively stable year to year

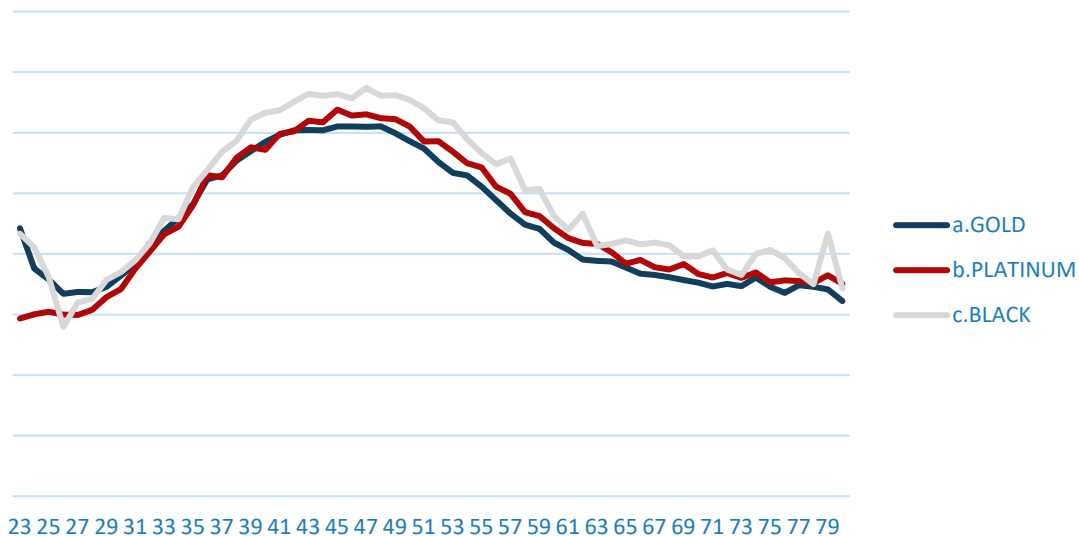
Cons

- **Doesn't capture all spend (multi-banked)**
 - **More cohort and period effects on the amount of spend as the percentage of consumption made through discovery may change for reasons other than year and age**
 - **Transactions may be misclassified (and hard to classify EFTs)**
-



Examples from the Data

Average family size by age 23-80



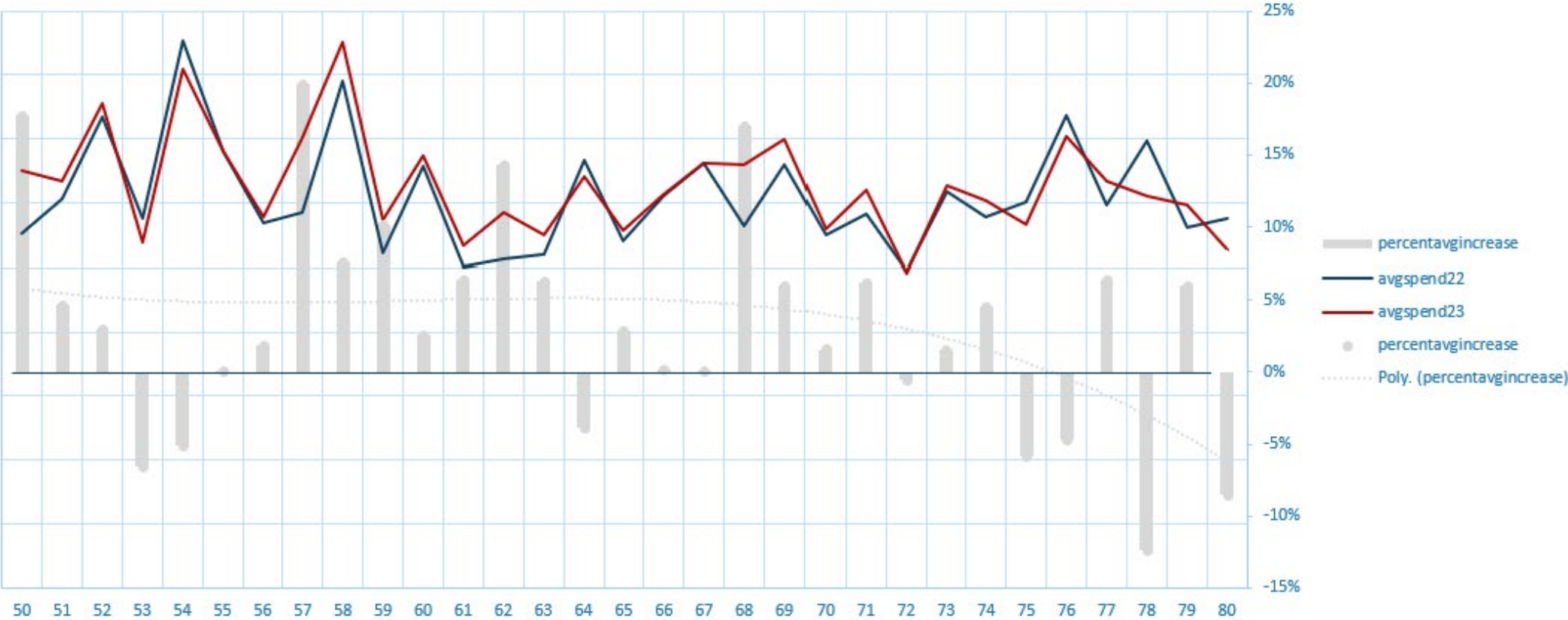
Sample of over 300,000 clients,
Over 100 at each age group up to age 83

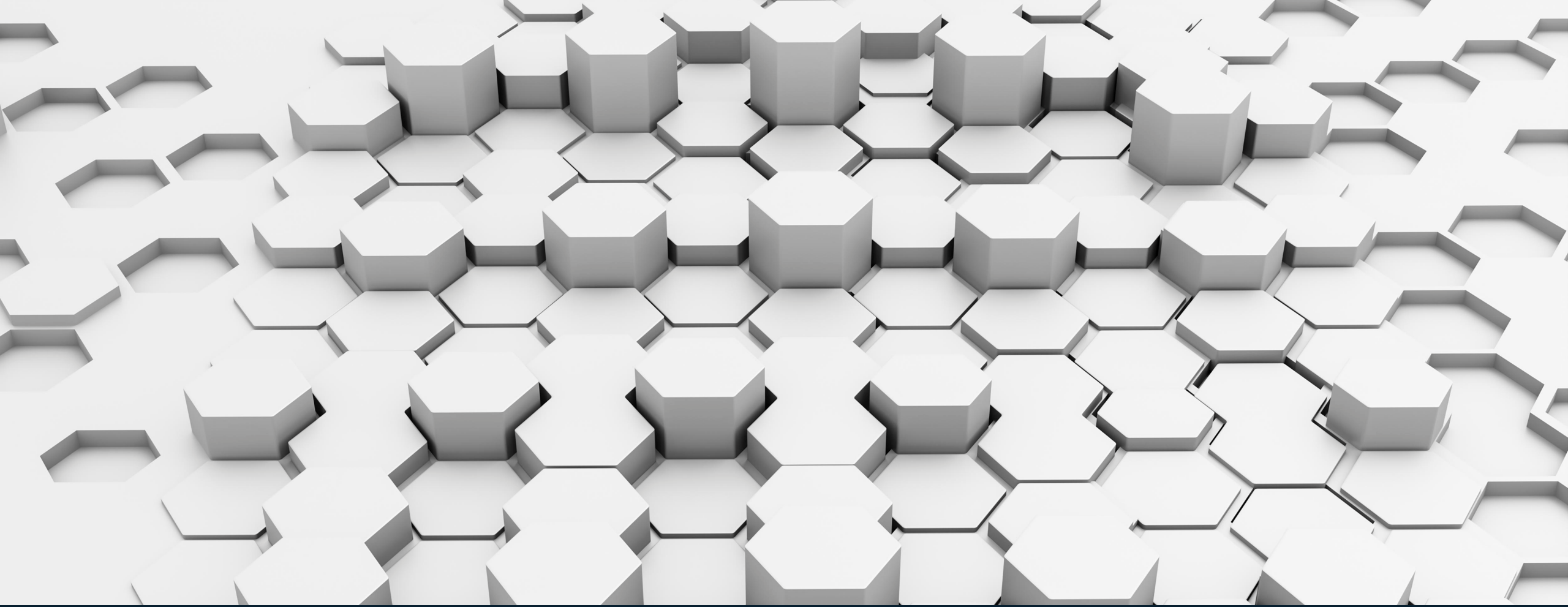
Age in 2022	# in 2022	# in 2023	Average Spend 2022 (rebased)	Average Spend 2023	Percent change from 2022 to 2023
60	3903	3898	1.00	1.15	15%
61	3798	3789	0.96	1.11	15%
62	3494	3494	0.97	1.13	16%
63	3131	3122	0.99	1.16	17%
64	2848	2844	1.04	1.14	9%
65	2457	2460	0.96	1.04	8%
66	2271	2275	1.00	1.11	11%
67	2003	2002	1.00	1.12	12%
68	1885	1884	0.90	1.08	20%
69	1619	1618	0.92	1.00	9%
70	1531	1529	0.90	0.95	6%
71	1180	1176	0.86	0.94	9%
72	1112	1107	0.84	0.89	6%
73	971	971	0.86	0.93	7%
74	858	851	0.83	0.90	8%
75	750	749	0.84	0.84	0%
76	566	563	0.90	0.91	1%
77	514	505	0.83	0.90	9%
78	448	445	0.96	0.92	-5%
79	326	323	0.74	0.83	12%
80	269	266	0.79	0.80	1%
81	218	216	0.79	0.76	-4%
82	157	155	0.81	0.85	4%
83	101	100	0.56	0.55	-2%



Examples from the Data

For each 2022 age cohort, that group's average spend in 2022 and one year later (when they are one year older)





Age, period and cohort effects



ACTUARIAL SOCIETY
OF SOUTH AFRICA

Advancing Actuarial Excellence

Educate • Develop • Uphold

Age, period and cohort effects



- Linear Dependence
 - Age = Period – Cohort
 - Example...
- Once you know 2, the last is follows
- Did we discover this?
- The official name is...
- The APC Identification problem
- Three observations have effectively 2 degrees of freedom
- They cannot be modelled separately

Social Science & Medicine 93 (2013) 163–165



ELSEVIER

Contents lists available at SciVerse ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed



Commentary

The impossibility of separating age, period and cohort effects



Andrew Bell^{a,b,*}, Kelvyn Jones^{a,b}

^a School of Geographical Sciences, University of Bristol, University Road, Bristol BS8 1SS, UK

^b Centre for Multilevel Modelling, University of Bristol, 2 Priory Road, Bristol BS8 1TX, UK

ARTICLE INFO

Article history:
Available online 4 May 2013

Keywords:
Age–period–cohort models
Model identification
Collinearity

ABSTRACT

This commentary discusses the age–period–cohort identification problem. It shows that, despite a plethora of proposed solutions in the literature, no model is able to solve the identification problem because the identification problem is inherent to the real-world processes being modelled. As such, we cast doubt on the conclusions of a number of papers, including one presented here (Page, Milner, Morrell, & Taylor, 2013). We conclude with some recommendations for those wanting to model age, period and cohort in a compelling way.

© 2013 Elsevier Ltd. All rights reserved.

Introduction

This commentary addresses this issue of statistically modelling separate age, period and cohort (APC) effects. The issue has been hotly debated for decades (Ryder, 1965), in sociology (Glenn, 1976; Mason, Mason, Winsborough, & Poole, 1973), medical science (Osmond & Gardner, 1989; Robertson & Boyle, 1986) and elsewhere. However, the publication of an article in this issue (Page et al., 2013) alongside more recent methodological developments in APC modelling (Tu, Smith, & Gilthorpe, 2011; Yang & Land, 2006; Yang, Schulhofer-Wohl, Fu, & Land, 2008), shows that there is still profound interest in modelling and discerning APC effects.

This commentary does not directly critique Page et al.'s paper in terms of its substantive conclusions; rather, we address the key methodological issues in modelling APC effects. However the implications of our argument call into question the results found by Page et al., and should act as a warning for others researchers wishing to disentangle APC effects in a meaningful way.

Age, period and cohort effects

The difference between age effects, period effects and cohort effects is well explicated by this fictional dialogue by Suzuki (2012, p. 452):

A: I can't seem to shake off this tired feeling. Guess I'm just getting old. [Age effect]

B: Do you think it's stress? Business is down this year, and you've let your fatigue build up. [Period effect]

A: Maybe. What about you?

B: Actually, I'm exhausted too! My body feels really heavy.

A: You're kidding. You're still young. I could work all day long when I was your age.

B: Oh, really?

A: Yeah, young people these days are quick to whine. We were not like that. [Cohort effect]



THREE COMPARISONS TO BE MADE

With data across ages and across time; we can make the comparisons

- Longitudinal comparisons
 - Same people, different times, different ages
 - Measures minimal cohort effects
 - Measures primarily age and period effects
 - This can show that 38 year olds of 2025 spend differently from when they were 21 years old in 2008
- Cross-sectional comparisons
 - Different people, same time and different ages
 - This can show that 21 year olds of 2025 spend differently compared to 38 year olds of 2025
- Time-lag comparisons
 - Not as useful for lifetime analysis, doesn't offer comparison across ages
 - Different cohorts/people, different times and the same age
 - Captures primarily the period and cohort effects
 - This could show that 21 year olds of 2025 spend differently than 21 year olds of 2008



OUR ANALYSIS

We observe people through 2022 and 2023

- Inspired by Macdonald & Moore's paper
 - All longitudinal inferences are driven by:
 - Age effects, this is what we hoped to observe in isolation
 - Period effects – the thorn in our side
 - Inflation between 2022 and 2023 are assumed to explain the year effects
 - The rest of the variance is attributed to the age effects
 - Several problems with this:
 - Covid spending recovery is NOT an age effect factor
 - Engagement with Discovery Bank, our data source, is not an age effect factor
 - Any changes in demographic mix of underlying data are not an age effect factor
-

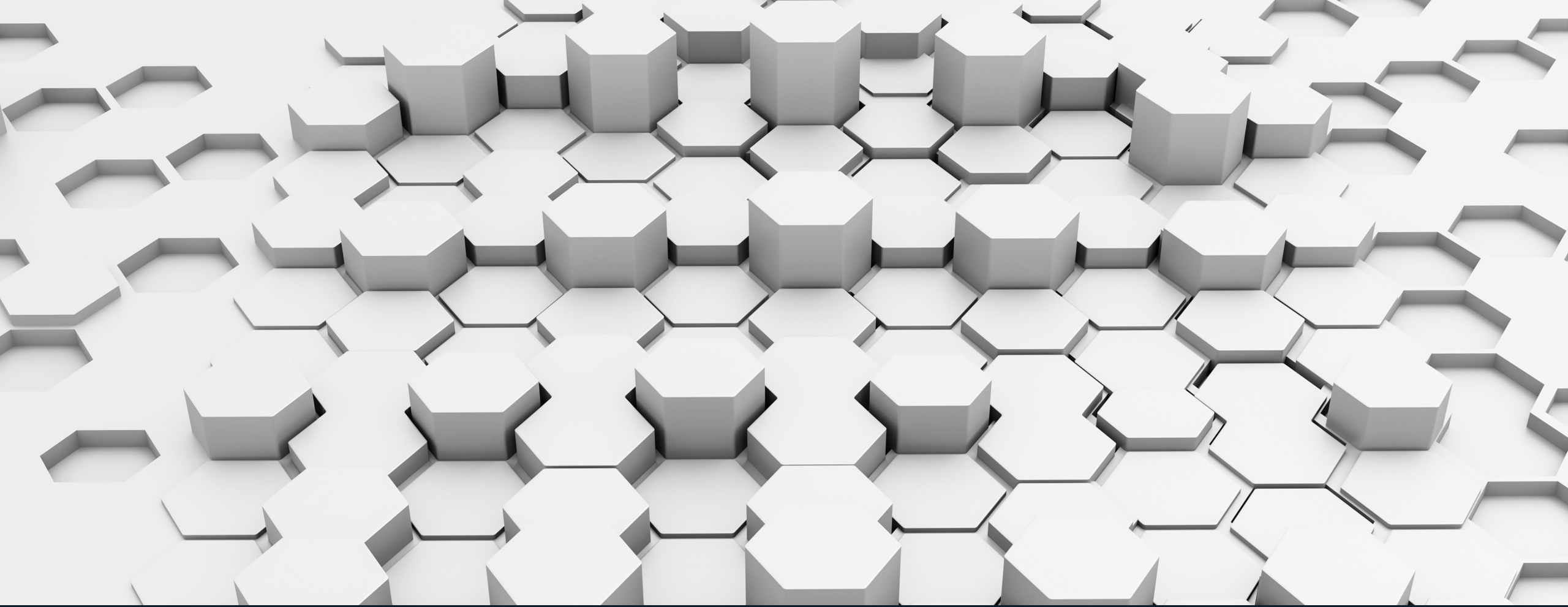


OUR ANALYSIS

Cross sectional observations

- Including the illustrations from Nkosinathi include age effects, the year is constant and contributes no effects
 - By their linear dependence, a cohort effect is induced
- This is much more onerous to quantify, however conventional wisdom tells us there's a difference
 - Boomers are often more responsible with investments, credit and home ownership
 - Whereas Gen Z is often riskier in their investments and in using credit
 - Bitcoins, Klarna, BNPL

This make it very difficult to make any clear observations with age as the single independent variable



Results



ACTUARIAL SOCIETY
OF SOUTH AFRICA

Advancing Actuarial Excellence

Educate • Develop • Uphold

RESEARCH QUESTIONS



- Is there a significant change in consumption at retirement relative to just before retirement?
 - How does consumption change over people's lives, both overall spend and per category, and particularly how does consumption change in retirement?
-

RESULTS



AVERAGE CONSUMPTION CHANGES IN 2023 MONEY TERMS

Category	Average Change (%)
All data	3.3
Vitality Money Gold+	7.81
Vitality Health Gold+	5.04
Gold Card Holders	1.14
Platinum Card Holders	4.86
Black Card Holders	2.16



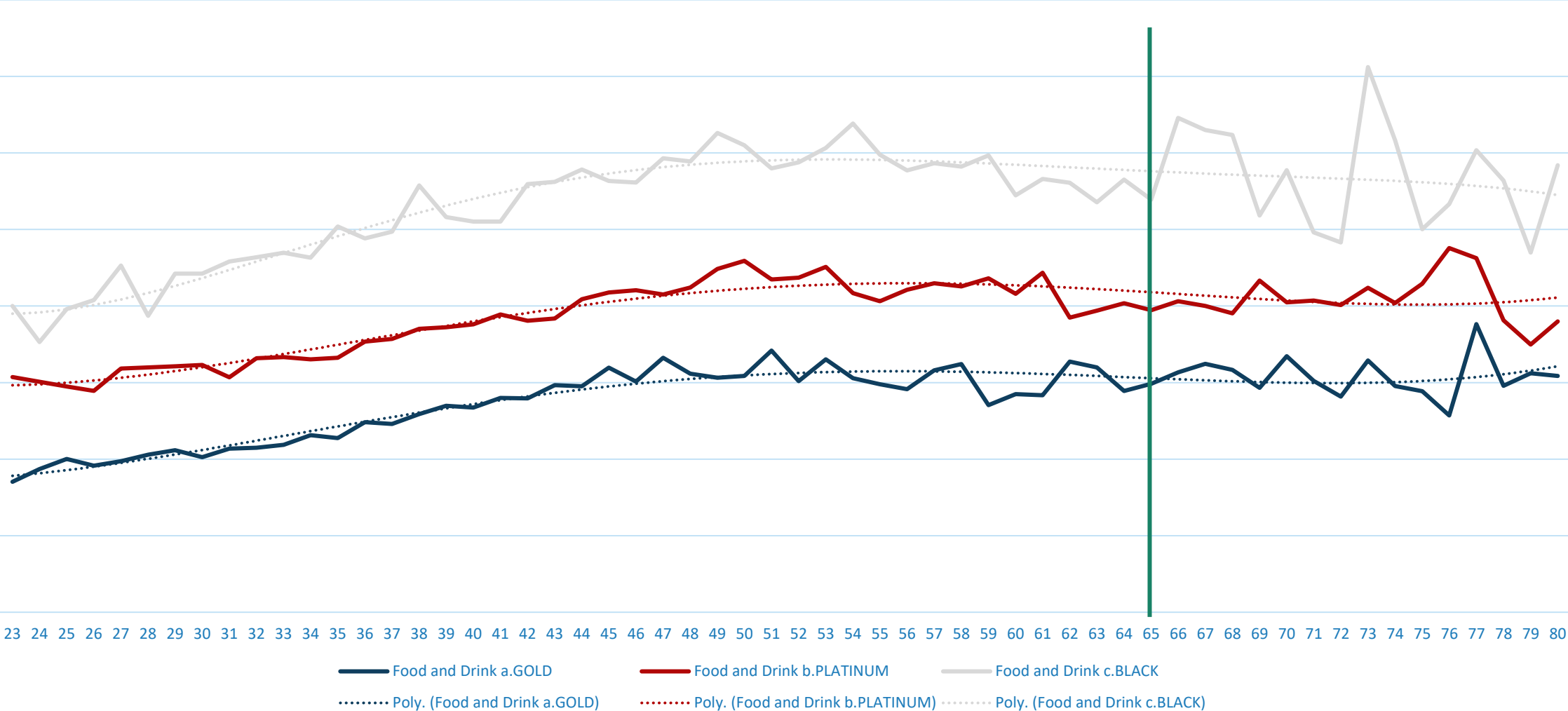
Hypothesis test

- μ_x : mean consumption at age x
 - $H_0: \mu_{64} \cdot 1.06 = \mu_{65}$
 - $H_1: \mu_{64} \cdot 1.06 \neq \mu_{65}$
 - Assumption: Independence
 - $\alpha = 0.05$; Non-rejection region: $[-1.96, 1.96]$
 - $T = \frac{\mu_{64} \cdot 1.06 - \mu_{65}}{S_p \sqrt{\frac{1}{n_{64}} + \frac{1}{n_{65}}}} = 0.4438$
 - $T \in [-1.96, 1.96] \gg$ Fail to reject H_0
-

RESULTS



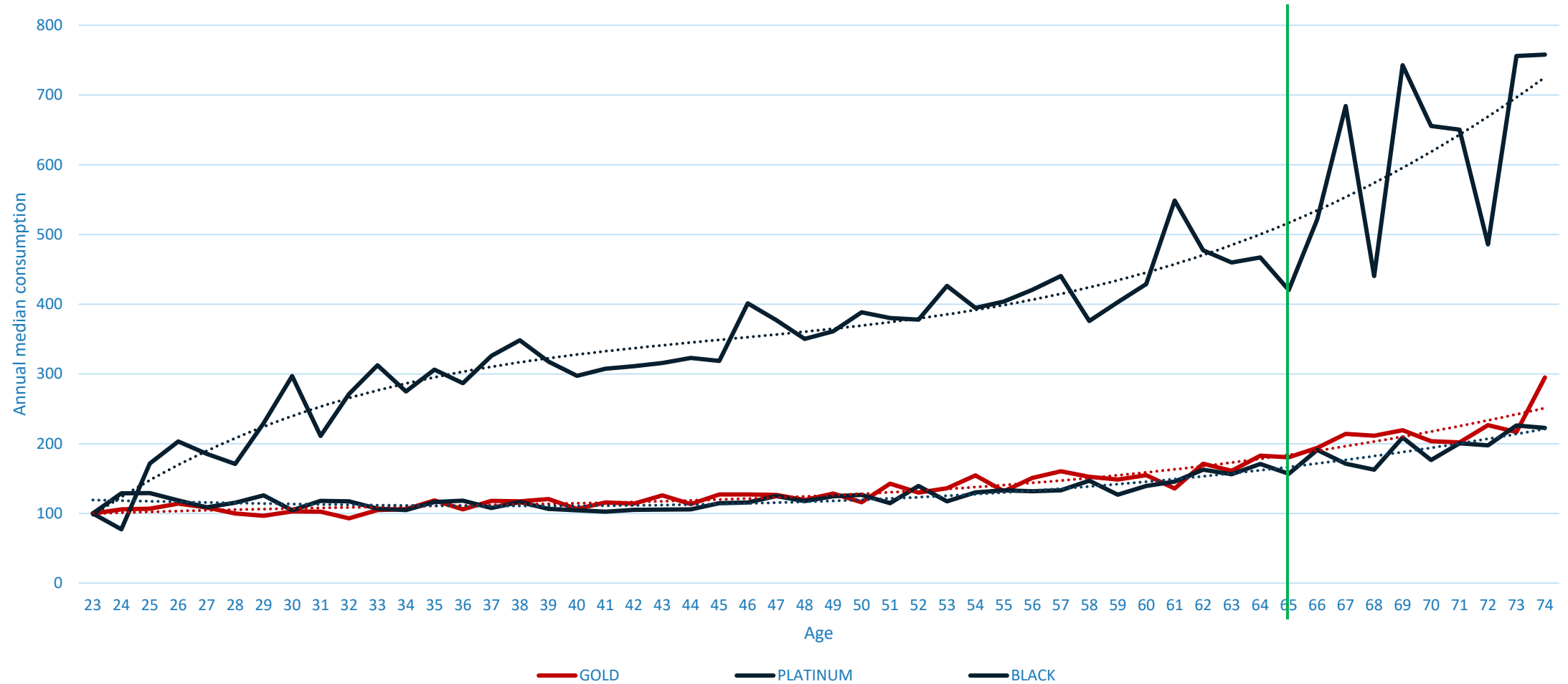
Trends in average spend on food and drink at different ages for different card colours



RESULTS



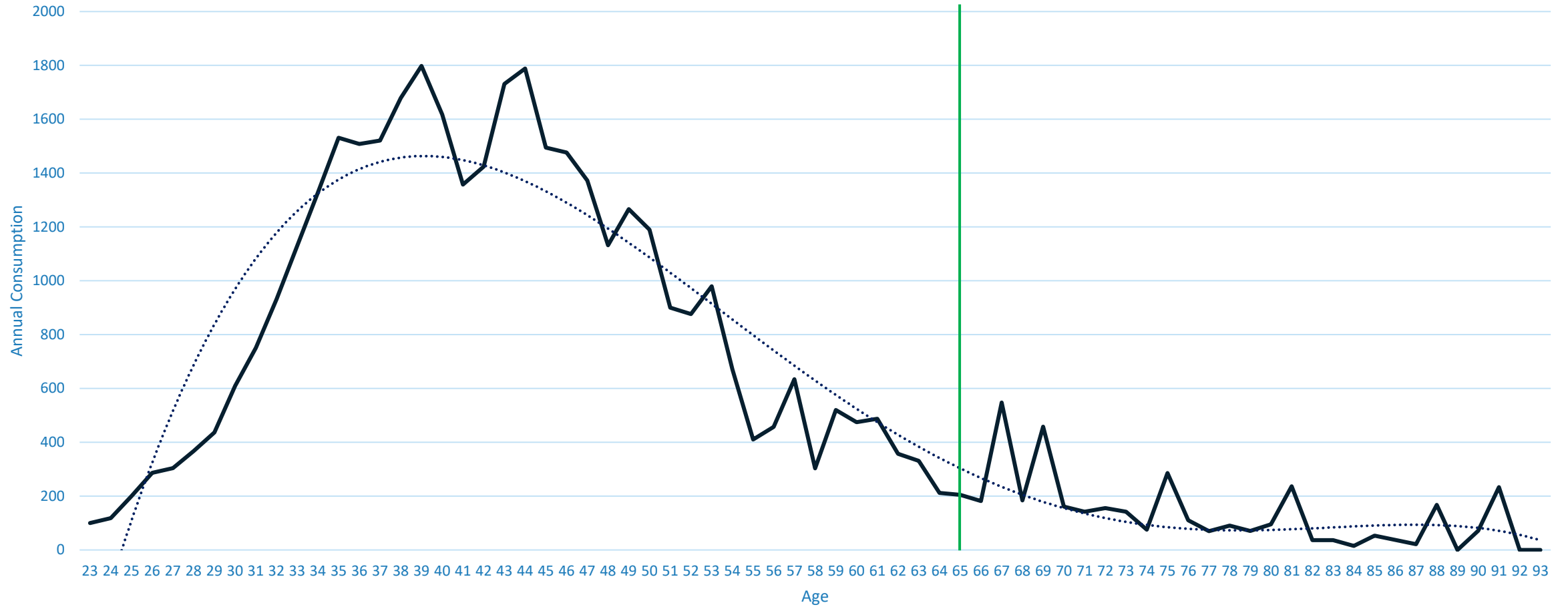
Health and personal care consumption across different ages and different card colours, rebased at R100



RESULTS



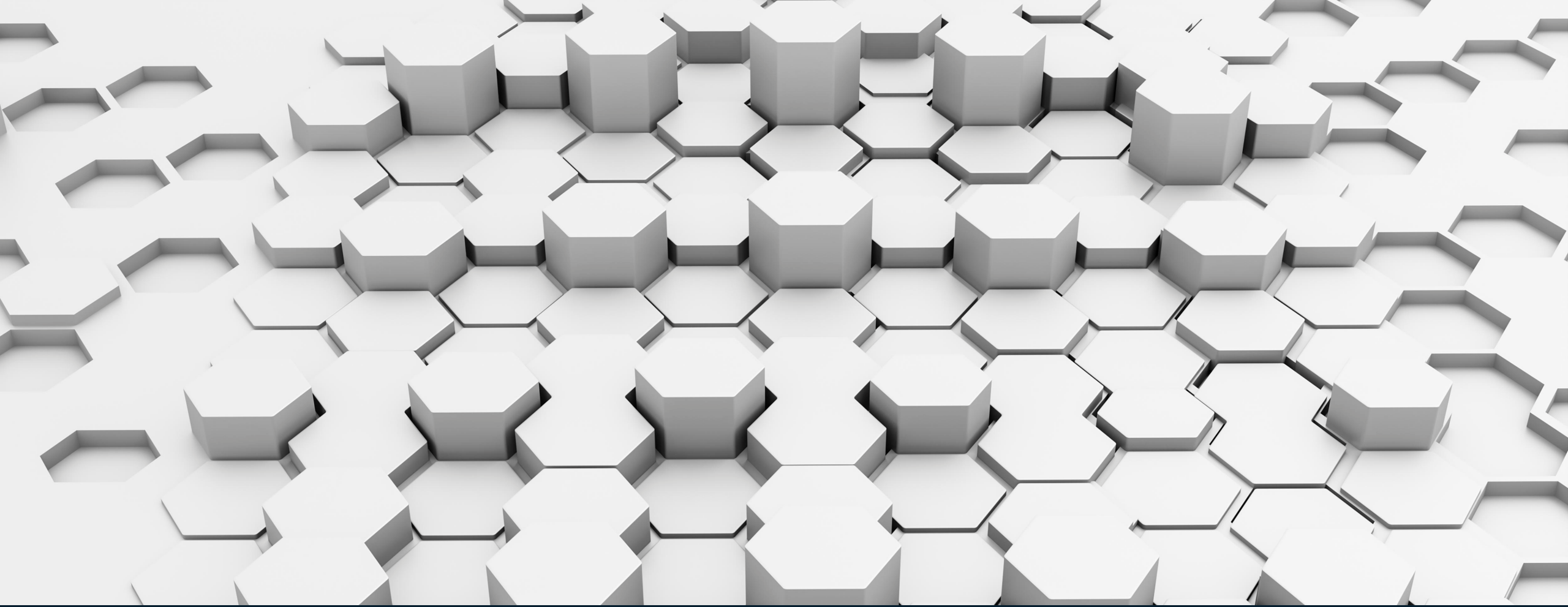
Consumption on children across different ages, rebased to R100 at age 23



CONCLUSION



- On average, consumption increases after retirement, although hypothesis test fails to conclude change in mean spend
 - Spending in categories like clothing and work decrease in retirement, whereas the more necessary categories of expenditure like food and healthcare are shown to increase in retirement
 - More research is required to confirm these macro findings, enable more personalized predictions, make use of more of the available information and to consider the appropriateness of a 100% consumption replacement ratio
-



THE END. QUESTIONS?



ACTUARIAL SOCIETY
OF SOUTH AFRICA

Advancing Actuarial Excellence

Educate • Develop • Uphold