[Page de garde]

HEC MONTRÉAL

Planning and Wealth Accumulation: Evidence from Canadian Survey Data

par

Joseph Rifaat

Decio Coviello
HEC Montréal
Directeur de recherche

Sciences de la gestion
(**Spécialisation Économie appliquée**)

Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences en gestion
(M. Sc.)

Décembre 2023

Résumé

Qu'entendons-nous par « richesse »? Comment la richesse devrait-elle être mesurée? La tendance à planifier joue-t-elle un rôle dans l'accumulation de la richesse chez les particuliers? À l'aide de données d'enquête recueillies auprès de plus de 40 000 Canadiens en 2017 et 2021, ce mémoire apporte deux contributions. Premièrement, un nouvel indice multidimensionnel est construit à l'aide de l'analyse en composantes principales, afin de mesurer la richesse individuelle. Deuxièmement, l'indice est utilisé pour étudier l'impact de la tendance à planifier sur la richesse, en utilisant une approche en variables instrumentales pour isoler l'effet causal de la planification. L'estimation de cet impact est difficile en raison de la possibilité des biais de simultanéité et de variable omise. Nous utilisons deux variables instrumentales pour la planification économique: l'adhésion à une routine et les capacités organisationnelles (des compétences non cognitives). Les résultats soutiennent la conclusion que la tendance à planifier entraîne l'accumulation de richesse. La relation entre la richesse et la planification est étudiée avant et après le début de la pandémie de la COVID-19 et les résultats suggèrent que la planification donne lieu à une plus grande résilience face au choc. Les estimations de régression quantile suggèrent que l'impact de la planification sur la richesse est plus élevé chez les individus les moins riches. Les résultats sont robustes aux différentes approches utilisées pour calculer l'indice de richesse, à plusieurs spécifications de modèles de régression, et ce, en utilisant des variables de contrôles pour d'autres déterminants de la richesse, notamment l'aversion au risque et la patience.

Mots clés : déterminants de la richesse; propension à planifier; planification économique; littératie financière; bien-être financier

Méthodes de recherche : analyse multivariée

Abstract

What does it mean to be wealthy? How should we measure wealth? How does planning enable individuals to accumulate wealth? Using survey data from over 40,000 Canadians collected in 2017 and 2021, this thesis makes two key contributions. Firstly, a novel multidimensional Wealth Index is developed using principal components analysis, to measure individual wealth. Secondly, the index is used to study the causal impact of individual planning on wealth, using an instrumental variable strategy to isolate the causal effect of planning. Estimating this impact is difficult because of the possibility of simultaneity and omitted variable bias. We use adherence to a routine and organizational skills (i.e., non-cognitive abilities) as instrumental variables for engagement in economic planning. The results support the conclusion that the propensity to plan causes wealth accumulation. The wealth-planning relationship is studied before and after the onset of the COVID-19 pandemic and the results indicate that planners showed more resilience in the face of the shock. Quantile regression estimates suggest that the impact of planning on wealth is significantly higher for lower-wealth individuals. Results are robust to different approaches used to compute the Wealth Index, to several regression model specifications, and when controlling for other determinants of wealth, notably risk aversion and patience.

Keywords: determinants of wealth; propensity to plan; economic planning; financial literacy; financial well-being

Research methods: multivariate analysis

Table of Contents

Résumé iii

Abstract iv

Table of Contents v

List of Tables and Figures vi

List of Abbreviations vii

Acknowledgments viii

1. Introduction 1

2. Context and Data 6

    2.1 The Canadian context 6

    2.2 The data 9

    2.3 Comments on data types and data quality 11

3. Literature Review 14

4. Measuring Wealth 25

    4.1 What is wealth? 25

    4.2 Methodology: the Wealth Index 28

    4.3 Results 33

5. Economic Planning and Wealth Accumulation 47

    5.1 Methodology: the wealth-planning relationship 47

    5.2 Results 53

    5.3 The propensity to plan and the pandemic shock 69

    5.4 Financial literacy 75

Discussion 78

6. Conclusion 89

References 90

Appendices i

    Appendix 1: Lifestyle Survey Privacy Notice i

    Appendix 2: Full PCA results iv

    Appendix 3: Alternative Wealth Index (‘shadow analysis’) xvi

    Appendix 4: Full quantile regression Results xxiii

    Appendix 5: Additional Regression Results xxv

    Appendix 6: Comparison of Wealth Index components in 2017 and 2021 xxxii

    Appendix 7: Data Definitions of Recurring Variables xxxiii

List of Tables and Figures

Table 1: Representativity checks for the 2017 and 2021 samples 10
Table 2: Dataset criteria for a repeated cross-sectional design 31
Table 3: Extract of PCA results from the global sample and qualitative analysis 33
Table 4: Factor loading scores for the retained variables 35
Table 5: Summary statistics for the Wealth Index and its component variables 36
Table 6: Evolution of the components of the Wealth Index between 2017 and 2021 37
Table 7: Selected differences between the core and alternative indexes 46
Table 8: Baseline OLS results for the effect of planning on wealth 54
Table 9: First-stage IV regression results 55
Table 10: Second-stage IV regression results (full sample) 58
Table 11: Robustness checks for planning and routine variable interactions 60
Table 12: Robustness check of the wealth-planning relationship 61
Table 13: Regression results with additional cognitive and non-cognitive factors 64
Table 14: QR results at each decile of the Wealth Index distribution 66
Table 15: QR results and implied percentage impacts using log of wealth 68
Table 16: IV regression results for 2017 and 2021 samples 69
Table 17: IV model, alternative specification with year dummy 71
Table 18: Evolution of mean and median Wealth Index between 2017 and 2021, by propensity to plan 73
Table 19: Modified IV regression models with planning, financial literacy, and interaction terms 76

Figure 1: Distribution of the Wealth Index variable 35
Figure 2: Age-Wealth Index profile for the global sample (N = 36,473) 38
Figure 3: Age-Wealth Index profiles for the global sample, by gender (N = 36,473) 39
Figure 4: Age-profile for each Wealth Index component 41
Figure 5: Cohort-Wealth Index profiles, 2017 sample and 2021 sample 43
Figure 6: Age-Wealth Index profiles, by highest level of education reached 44
Figure 7: IV regression schema 50
Figure 8: Mean propensity to plan, by age 62

List of Abbreviations

CMHC	Canada Mortgage and Housing Corporation
DHS	Demographic and Health Surveys
FHSA	First Home Savings Account
GDP	Gross Domestic Product
IV	Instrumental Variable
OCEAN	Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary Least Squares
PC	Principal Component
PCA	Principal Component Analysis
QR	Quantile Regression
TFSA	Tax Free Savings Account
USAID	United States Agency for International Development
WI	Wealth Index

Acknowledgments

First and foremost, I'd like to express my gratitude to my research director, Decio, for making this work possible. I cannot say how thankful I am to you for facilitating my use of this dataset and for providing so much invaluable advice and feedback throughout my thesis. Thank you for helping me to navigate the intricacies of this subject that I am passionate about.

David, I am immensely grateful for the trust you put in us with the survey data. Thanks to you, Walter, Laurianne, and Honey for taking the time to share your comments on the initial results, and for our fruitful discussions on the topic of individual financial well-being.

Thank you to Jocelyne and Jean C. Monty; I appreciate your generous support through your contribution to the finance scholarship that I received as I started work on this thesis.

Support from my friends, family, and colleagues has been instrumental in carrying out this research. In particular, thanks go to Sonny for the review of my draft and for his insights, to Anne Sophie for the reading recommendations, and to my mother, Jo, for the moral support throughout my master's.

Special thanks to my boyfriend, Mitch, for your patient encouragement and for getting me over the finish line in the final few weeks.

1. Introduction

The concept of wealth is both omnipresent and contentious in economics[1]. 'Wealth' can be used in a strictly mathematical sense to mean 'net worth', in a more holistic sense to include the notion of human capital, or in a semantic sense to be synonymous with 'prosperous' or 'affluent'. Both the definition and measurement of wealth have always presented challenges to economists, who are invested in understanding what determines wealth and why it is unevenly distributed among individuals. Using unique survey data obtained from a sample of over 40,000 Canadians in 2017 and 2021, this thesis makes two core contributions. Firstly, it proposes, builds, and evaluates a novel indicator to measure individual wealth. The 'Wealth Index' is constructed through principal components analysis (PCA) from data obtained through a survey of individuals' socioeconomic, demographical, and psychometric attributes. By creating an index, this technique enables us to establish a measure for wealth that does not rely on net worth data, on which reliable data is difficult to obtain. Secondly, this index is used to estimate the causal relationship between individual economic planning and wealth accumulation, using an instrumental variable (IV) strategy. The causal effect of planning is estimated while accounting for key determinants of wealth from the literature, such as risk and patience. Economic planning by individuals is the central topic of this objective. The terms 'planning' and 'propensity to plan' are used in this thesis as being synonymous to economic planning. Where a different type of planning is meant, it is described as such, e.g., 'financial planning' or 'non-economic planning'.

Estimating the causal impact of planning on wealth presents a statistical challenge. An individual might engage in planning because they become wealthy, and not the other way around. Conversely, a low net worth individual might feel that planning is not useful, because they are resigned to not accumulating wealth. In these cases, the issues of reverse causality (the effect of wealth on planning) and omitted variables (the omission of

[1] The contentiousness of wealth will be explored in Sections 3 and 4.1. Key debates include: what assets are included in the definition (e.g., Advani et al., 2020); how wealth data can be obtained in a context of poor general understanding (Lusardi et al., 2017); whether wealth should include intangible elements such as education and health (Arrow et al., 2012).

'fatalistic attitude') would cause estimates of the causal impact to be biased. An instrumental variable strategy is conceived in light of these issues and is presented in detail in Section 5.1. Based on the variables available in the dataset, two instruments are used: the tendency to follow a routine and the tendency to make plans for oneself and one's friends. Although no instrumental variable can be 'perfect', these instruments were chosen on the grounds that they performed well in tests of the two key IV criteria (relevance and exclusion restriction). They represent non-cognitive skills that have a significant positive impact on economic planning, but are assessed to have a strong direct influence on wealth accumulation itself. In other words, their only relationship with wealth is through economic planning. The IV strategy is also used to specify two subsequent models to address the research questions. Firstly, a quantile regression (QR) design is used to evaluate the causal impact of planning on wealth at different parts of the wealth distribution. Secondly, an analysis of the effects of the COVID-19 shock on the wealth-planning relationship is also performed to take advantage of the samples from before and after the onset of the pandemic.

Three objectives underly the research undertaken for this thesis:

1. Build a novel indicator of individual wealth;
2. Evaluate the causal impact of economic planning on wealth accumulation;
3. Assess the relationship in the context of the shock of the COVID-19 pandemic.

The first objective is the construction of a quantitative index that measures several dimensions of wealth. Following examples of methodologies used to generate comparable indexes (described in more detail in Section 4.2[2]), the Wealth Index will be constructed through the execution of principal components analysis (PCA). This statistical technique, used for dimensionality reduction and index construction, is used on over 400 variables from the survey data in order to identify which variables are suitable for inclusion in the index calculation and how each of them should be weighted. Four variables are retained to build the index for each individual: household income, homeownership, self-assessed

[2] This thesis will use Filmer & Pritchett's (2001) approach to constructing a wealth index in India, which has been used and modified by USAID's Demographic and Health Surveys (DHS) program. Vyas and Kumaranayake (2006) provide an overview of the methodology and its use in this context.

financial situation, and financial anxiety. It is difficult to ascertain the accuracy of the Wealth Index, because no objective 'true' value exists as a comparison point. That said, the index performed well when used to construct age-wealth profiles: results were consistent with theoretical predictions related to age, gender, and education levels. As for the selection of theories analyzed in the literature review, the criteria were was follows: relevance to the survey data (theories were retained only if they covered the same subject areas as the data), clear theoretical evolution (widely discredited models were excluded) and repeatability (existing studies were favoured if they could be reproduced and retested with the dataset). As described in Section 4.3, two different Wealth Index indicators are constructed for the purposes of this research. One is used in the main part of the thesis, while the other is used to replicate all of the models and can be found in Appendix 3. This has two aims. Firstly, it provides a robustness check to the results of the main Wealth Index and enables us to have greater confidence in the performance of the index. Secondly, it seeks to show that methodologies used to create a more holistic measurement of wealth are inherently subjective and can be constructed in different ways.

We use this index to causally estimate the role of individuals' economic planning behaviour on wealth accumulation. This builds on the notion of the 'propensity to plan', which emerged as a determinant of wealth in the early 2000s and has been the subject of an increasing amount of research in the field of behavioural economics. Several empirical studies have examined the role of financial planning in wealth accumulation (e.g., use of financial instruments, development of a financial plan). This research will focus on economic planning, a broader category, that includes financial planning but also covers all of the ways in which individuals use their resources. In this sense, economic planning, even on an individual level, is analogous to large-scale plans to decide how to allocate resources in order to achieve prosperity or success. With these definitions, someone who carefully considers and compares options before making a consumption purchase is engaging in economic planning, but not financial planning. Similarly, an individual who decides to invest in obtaining an additional level of education in order to improve their wealth accumulation potential is engaging in economic planning, but not financial planning. Conversely, someone who regularly consults with a financial planner and has long-term financial goals and strategy is engaging in both financial and economic

planning. The results indicate a strong, significant relationship: a one-standard deviation increase in planning corresponds to a 9.3% increase in wealth. The estimation is robust to a several different specifications, two measures of planning, and both approaches used to construct the Wealth Index. Our models also include a control variable for ethnicity/race and for residents of rural areas, both of which are key sociodemographic determinants of wealth in the Canadian context (and indeed in other countries). These covariates show a strong positive relationship between wealth levels for white respondents (the 'racial wealth gap') as well as those living in cities (the 'rural-urban wealth disparity') which gives support to existing empirical evidence and shows the importance of including these control variables in our model of wealth accumulation.

By providing separate model results for each decile of wealth, quantile regression shows how the causal impact of planning on wealth varies in individuals of different levels of wealth, which provides two valuable insights. Firstly, it identifies whether the relationship is generalized across the population or if it is confined to specific sub-groups. Secondly, it provides more precise implications of the results by highlighting groups where the relationship is stronger and which individuals might see greater wealth gains from an increase in their planning activities. The results indicate that the wealth-planning relationship is strong and significant at all deciles of wealth, implying that the effect is general and not local to one specific level of wealth. This result is robust to two different specifications and with both the 'core' and 'alternative' Wealth Index variables. Furthermore, QR estimates provide evidence for a novel finding that the wealth-planning relationship is stronger in lower-wealth individuals relative to their levels of wealth. By one specification, the relationship is over twice as strong amongst the lowest decile of wealth than the highest.

The third objective takes advantage of the fact that the data was collected before and after the onset of the pandemic in order to examine the wealth-planning relationship in the context of the economic shock of COVID-19. The shock lends itself well to this analysis due to its wholly unpredictable nature (we can reasonably assume that nobody had prior knowledge of the shock) as well as the large scale of the event (the entire Canadian population was affected, albeit in different ways). Given that the propensity to plan speaks

to individuals' resilience in the face of uncertainty, the relationship is of particular economic interest, and results in a relevant and novel contribution to the empirical evidence on the role of planning in determining wealth. The results, obtained through repeated cross-sectional analysis, indicate that individuals with a higher propensity to plan experienced better wealth outcomes following the onset of the pandemic, even when controlling for key socioeconomic attributes such as education, marital status, race/ethnicity and living in a rural/urban area.

The structure of this thesis broadly follows the objectives described in this section. Section 2 provides background information on wealth in Canada as well as the survey data used in this research. Section 3 identifies and describes key casual mechanisms of wealth accumulation and introduce testable hypotheses for analysis. Section 4 is divided into two parts: the first outlines the methodology followed to build the novel indicator, while the second examines how the indicator performs when the hypotheses from the literature are tested. Section 5 describes the approach to evaluate the relationship between planning and wealth accumulation, followed by the results of the regression analysis and the comparison of those results between pre- and post-pandemic samples. Section 6 critically analyses the results and limitations, and presents implications for individuals, organisations, and policymakers. Section 7 provides a brief conclusion and a summary of the contributions of this thesis, along with future avenues for research.

2. Context and Data

This section presents a brief portrait of individual wealth in Canada and introduces some of the issues that specifically affect Canadians: precarity, debt and financial literacy. The unique dataset, the Lifestyle Survey, is then introduced, along with commentary on its reliability and its appropriateness for the research objectives.

2.1 The Canadian context

Looking at individual affluence in its net worth sense, 'traditional' measures of wealth paint a rosy statistical portrait of Canada. Per capita Gross Domestic Product (GDP) is 5.7% higher than the Organisation for Economic Co-operation and Development (OECD) average (OECD, 2022); the country is considered broadly as a 'rich country among rich countries'. Statistics Canada (2022) reported that median household net worth was $950,558 in the second quarter of 2022, which is very high by global standards. Wealth inequality exists and persists in Canada but is lower than in the United States: there, the wealthiest 1% hold approximately 31.3% of wealth, compared with 24.8% in Canada (Federal Reserve, 2021; OPBO, 2021). However, Canadians' financial well-being is negatively impacted by several well-known structural problems.

As it relates to savings and budgeting, between a third and half of Canadians do not have sufficient funds to cover emergency expenses of over $2,500 (CPA Canada, 2022); this figure is just 4% of median after-tax annual income (Statistics Canada, 2023). The onset of the pandemic, with widespread furloughs, job losses and business closures, revealed the precarity of individuals' financial situations. While federal and provincial governments acted swiftly and impactfully to provide short-term relief, their policies were financially costly and may have led to adverse economic effects in the medium-term.

Household debt is another well-documented systemic issue in Canada. Contrary to the country's performance on other indicators, Canadians have the OECD's 9th highest level of household debt, which stands at 187% of household income, or 107% of national GDP (CMHC, 2023). While much of this burden is manageable, secured debt, a sizable portion of it is not. Particularly in the current context of rapidly rising interest rates, the Canada

Mortgage and Housing Corporation (CMHC) warns that the high debt burden increases Canadians' vulnerability in the event of a recession or other economic shock. Unlike the United States and the United Kingdom, which saw their individual debt burdens decline between the end of the Great Recession and the onset of the pandemic, Canadian household debt as a percentage of GDP rose by 12 percentage points (CMHC, 2023). This is particularly problematic in the context of access to homeownership, which remains the principal source of Canadians' household wealth (Harding and Rosenthal, 2017). Individuals with high personal or unsecured debt are disadvantaged if they attempt to enter the property market, and, especially in a context of high interest rates, even existing homeowners might become less able to contribute to equity on their own properties.

As is the case in many economies, financial well-being and literacy are unevenly spread across Canadians. Three socioeconomic groups are notably susceptible: women, low-income individuals, and Indigenous people (Lusardi & Mitchell, 2008; Financial Consumer Agency of Canada, 2021). While the complex history of this inequality will not be explored as part of this thesis, the results will be interpreted with a particular focus on identifying uneven impacts of planning on financial well-being. In response to these issues, the Canadian federal government and private Canadian institutions have put heavy emphasis on the promotion of financial well-being among the general population, including planning and literacy. Every five years, a National Financial Literacy Strategy is published to this end by the federal government, with the aim of creating "a Canada where everyone can build financial resilience in an increasingly digital world" (Government of Canada, 2021). New financial services industry players in Canada, such as *Wealthsimple* and *KOHO*, promote financial literacy and planning tools as part of their service offerings. Traditional banks, on observing the success of these strategies, are playing catch up.

As economists identify ways of quantifying wealth purely in terms other than net worth (e.g., comprehensive wealth, human capital, etc.), these underlying issues are receiving much-needed attention. Our definition of individual wealth is expanding from a simple formula of assets less liabilities to a more diverse series of perspectives that include financial security, well-being, literacy, and others. It is logical to assert that our measures

of individual wealth should evolve at the same time. This requires high-quality, standardized data that can be used to build a more complex statistical view of wealth.

2.2 The data

The dataset used in this research is from a survey conducted by a marketing insights firm in both English and French, across Canada. Two samples are used from 2017 and 2021 in which 22,274 and 18,328 unique respondents (for a total of 40,602 observations) provided information on a range of subjects, including: attitudes, demographics, education, income, homeownership, psychological profile, habits, and other data. The resulting dataset has over 3,000 variables across these categories. A similar methodology is also used by the firm and its subsidiaries in China, France, Germany, Italy, Japan, the Netherlands, and Spain. The data collection is a repeated random cross-section methodology: at each administration of the survey, a representative sample is selected, with no intentional overlap between respondents. If an individual were to complete the survey in more than one year, it would be coincidental and unintentional, and no identifier is available to track them longitudinally. The data is built to be representative of the Canadian population, particularly as it relates to socioeconomic status, geography, and immigration status. No identifiable information is contained within the dataset, individuals are anonymous and a randomly generated unique reference code is used as an identifier.

An initial verification was completed to ensure the representativity of the data population to the Canadian population; the results are summarized in Table 1. Key variables of interest show an encouraging degree of representativity. As it relates to income, a vital dimension of this research, the dataset is highly representative of the Canadian population. Some values for income are missing, from respondents who chose not to answer, but the proportion is small. Education levels are very representative when measured by the proportion of school leavers, though we observe an oversampling of university degree holders in 2017. Employment is well estimated in both samples and appears to capture the drop in employment by 0.2 percentage points before and after the pandemic. Homeownership is the variable with the highest difference from the true Canadian figure, with a gap of -6 percentage points in 2017 and -13 points in 2021. Though this is undesirable, the survey is internally consistent in its under-sampling of homeowners, and participants in surveys such as are these typically from lower-income or lower wealth population groups.

Table 1: Representativity checks for the 2017 and 2021 samples

	Observations (N)	Sample mean	Expected (population) mean
2017 sample (pre-pandemic onset)			
Annual household income	19,631	$46,024.27	$46,210.70
Education level, % not completed high school	21,998	9.0%	9.0%
Education level, % with bachelor's or higher	21,998	36.0%	31.0%
Employment status, % working full/part time	22,274	63.2%	66.5%
Homeownership, % owning home	22,274	61.7%	67.8%
2021 sample (post-pandemic onset)			
Annual household income	16,842	$53,568.46	$54,000.00
Education level, % not completed high school	18,328	7.0%	7.0%
Education level, % with bachelor's or higher	18,328	36.2%	35.0%
Employment status, % working full/part time	18,328	63.0%	66.3%
Homeownership, % owning home	18,328	53.1%	66.5%

The expected means were derived from Statistics Canada data on each of the five indicators from the following sources: Table 11-10-0239-01 (income), Table: 37-10-0130-01 (education), Census of Population (homeownership). For income, which is given in 2021 constant dollars, the Bank of Canada's inflation tool was used to calculate the equivalent in 2017 dollars. Given that the dataset does not contain data on wealth in a 'net worth' sense, no representativity comparison is made.

2.3 Comments on data types and data quality

Several characteristics of this dataset make it a uniquely rich source of information about individuals' wealth accumulation. Firstly, the combination of variables lends itself particularly well to identifying determinants of wealth from several types of indicators. As shown in Section 3, theories of wealth accumulation span cover demographic (age, marital status, geography) financial (income, parents' income), and human capital/cognitive (education, financial planning, financial literacy) themes. Detailed questions in the Lifestyle Survey generate precise and comprehensive data about habits, behaviours, attitudes and propensions that are not often available in similar large-scale data collection exercises, such as censuses. The result is that more nuanced conclusions can be drawn from the correspondence of information on financial well-being and determinants of wealth than in other similar studies. Secondly, given that the same methodology has been deployed in eight countries and over several years, the data produced is highly reliable. The survey is repeated regularly and contains questions that are adapted to each country's economic realities. Finally, the sample size of approximately 20,000 annual respondents aged between 15 and 75 is sufficiently large to obtain small confidence intervals and greater precision in the results.

As a proportion of an entire population, this can be considered to be 'nationally representative' (Dawson, 2023). This level of coverage is extremely high for private sector data and gives additional support to the conclusions that can be drawn from the analyses performed. Secondly, the wide scope of the questionnaire – generating a total of 5,000 variables – provides extremely rich information on each individual as well as opportunities to perform robustness checks that are not possible with other datasets.

The data is collected through a computerized survey completed by individual respondents, in their chosen language: within Canada, English and French are offered. Participants provide responses anonymously, and certain questions deemed particularly sensitive are optional (e.g., those related to religious beliefs, sexual orientation, and other personal aspects). The Privacy Notice is provided in full in Appendix 1. The fact that all the information within the survey is self-reported has both advantages and drawbacks. For certain variables, such as income or education, this is not a particular source of concern:

Micklewright and Schnepf (2007), for example, show that self-reported income data was very similar to official statistics in the UK. This is especially true when salary scales are given as options (although this results in some data loss) and when both individual and household data variables are collected. When it comes to more subjective questions, a survey is a highly efficient way of obtaining behavioural data such as habits and level of discipline, because these cannot be measured by objective scales except for in highly controlled circumstances. However, social desirability bias is problematic in these contexts because it is well known that respondents are much more likely to report answers that are seen as 'better' or closer to a perceived 'correct' point. Given the anonymous nature of this survey, there is no clear incentive for interviewees to knowingly provide false information, but the idea of bias is still considered in the analysis of the results.

Most of the behavioural questions are answered using a Likert scale from 1-5: strongly agree (5), agree (4), neither agree nor disagree (3), disagree (2), and strongly disagree (1). This type of data is useful for statistical analyses because it converts thousands of qualitative data points into interpretable units of information. We can use more objective questions to compare responses and check for internal validity and coherency. For example, we would expect to see a very high negative correlation between the age variable and the variable of time spent using a smartphone. A key advantage of Likert data is that the scale is extremely simple to understand and remains very similar across a variety of different questions, so the risk of respondents misunderstanding or misreporting data is comparatively low. However, certain issues arise from using Likert data. Comparability is one such problem: is it true that one respondent's 'agree' is the same as another's? If asked the question 'I have more ability than most people', two people with greatly different levels of real 'ability' would both conceivably reply 'agree' if they truly felt that they had a 'good' or above-average level of ability. As there is no way of adjusting data to an objective value, we must accept an assumption of comparability. Another issue stems from the mathematical properties of Likert answers. It is debatable that an 'agree' (i.e., 4 points) is worth 'double' that of a disagree (i.e., 2 points) and that a 'strongly agree' is worth five times that of a 'strongly disagree'. Additionally, the presence of a middle value (in this case, 3 'neither agree nor disagree), is sometimes used as a proxy for 'I do not know' or 'I do not wish to respond', even subconsciously. This is problematic from

both a mathematical perspective, in that ‘3’ ceases to have its logical value between 2 and 4, and from an interpretation perspective, because it is unclear if the respondent really means ‘neither agree nor disagree’, ‘I don’t know’, or ‘I don’t want to answer’. As a solution to these numerical problems, the example of the statement above would be statistically considered as asking ‘On a scale of 1 to 5, do you rate your ability as much better than most people (5) or much worse (1)?’. This is a pragmatic and workable approach to this dataset, though it is not taken for granted. Robustness checks to avoid interpretation problems are described and applied in response to this point.

3. Literature Review

This literature review will provide a critical overview of selected theories of wealth accumulation, starting with the neoclassical lifecycle model approach and finishing with contributions from behavioural economics. Recently, economists have called into question the 'net worth' definition and complemented it with more nuanced ways of conceptualizing and measuring wealth. From each theoretical 'family', hypotheses will be established such that they can be tested with the dataset, with a view to providing novel insights on the determinants of wealth.

Neo-classical theory: accumulating wealth over the lifecycle

Neo-classical lifecycle theory is the cornerstone of modern economical thought on wealth accumulation, and has been highly influential in shaping university curricula, financial institutions' strategies, and monetary policy models. Indeed, Bütler (2001) describes is as as 'textbook example', while Karlan and Morduch (2010) argue that it is still the most influential model of savings in economics. Given that the model originated in the 1960s, the 'net worth' definition of wealth applies throughout this sub-section.

As the name suggests, lifecycle theory looks at the change in an individual's affluence over time and predicts differing rates of accumulation and disaccumulation of wealth. The basis of the core model is based on a simple set of variables: labour income (y), consumption (c), and saving (s). At the start of an individual's working life, they begin to receive an income and will use the vast majority of that to consume goods and services. The savings rate – defined as $\frac{s}{y}$ – starts off low. Generally speaking, income rises over time: as an individual becomes more experienced, skilled, or specialized (or a combination), their labour becomes more valuable. While consumption rises too, the lifecycle hypothesis predicts that it rises at a much slower rate, which necessarily increases the proportion of income dedicated to saving (Ando & Modigliani, 1963). The reason for this is consumption smoothing, which is explained in the following paragraph. With an increasing amount of savings, the individual is able to accumulate wealth, for example in the form of liquid assets or property. With financial instruments, interest

earned compounds over time, which has an exponential impact on the individual's wealth. The same is true of property, whose value tends to increase at a rate faster than that of the economy overall. Empirical models tend to put the peak of labour income at between 50 and 55 (e.g., Bütler, 2001 and Scholz et al, 2006), which means that wealth tends to peak just before retirement, which is broadly defined as 60 to 65 years old. From retirement onwards, a period of disaccumulation ensues, where an individual uses their savings to consume and makes no more labour income. Plotted on an age-Wealth Index profile, the wealth curve is an asymmetrical hump-shape: it rises at an ever-increasing rate from the start of working life through to retirement, before declining into the later years of life.

The lifecycle model is built on several key assumptions. The first is consumption smoothing, a theoretical prediction that individuals make efforts to maintain a similar level of consumption throughout their life. This is observed due to an individual's expectations about their future (permanent) income, which is the revenue profile that is expected over the course of a working life. Taking an oversimplified example, if a recent graduate who earns $80,000 expected to steadily increase their earnings up to a lifetime maximum of $120,000, they might choose to smooth their consumption over their whole working life *as though* they had a constant income of $100,000. The individual incurs debt (i.e., a negative savings rate) in their younger years to fund consumption, and then as their income approaches the maximum, their savings rate increases and wealth accumulation peaks. Secondly, a hypothesis is established concerning the discount factor, which measures the patience of the individual: how much do they value one 'unit' of consumption today compared with consumption at a future date. Intuitively, high patience (i.e., attributing a higher value to future consumption relative to present consumption) is associated with greater saving and therefore greater wealth accumulation, while individuals who consume more of their income save less and accumulate less wealth. The discount factor is now understood more broadly to encompass not just patience itself, but also the ability and discipline to manage consumption and saving.

The lifecycle model has widely come to be understood as empirical fact: seven decades' worth of data have confirmed the general patterns that the core hypothesis predicts. The nuances are within the assumptions. A key criticism of this model was the idea of

consumption smoothing: how often do people meet individuals who have near-perfect certainty of their future earnings and who consume the same amount of goods and services throughout their entire working life? To do so would require a rather unrealistic level of confidence and accounting ability. Hall (1978) refined the model by introducing a random walk that accounts for income shocks (e.g., a promotion or a job loss) that would cause someone to abruptly change their level of consumption. Skinner (2007) echoes this point about individual shocks, as well as macroeconomic events such as stock market crashes, that can hinder smoothing. Furthermore, he raises the question of "obsessive oversaving", and whether smoothing is even optimal from a utility perspective. In this sense, an individual who dedicates too much time or effort to planning and saving might subtract more from their ability to enjoy their working years than they gain from an optimized financial plan. Consumption smoothing is also challenged by the liquidity restraint, whereby some individuals cannot gain access to credit to fund consumption above their current income level. Additionally, Pozzi and Sadaba (2023) demonstrate that consumption smoothing tends to hold less during economic shocks such as pandemics and wars, when expectations about future earnings are upended. In these circumstances, individuals tend to save and spend by 'rule of thumb' and think only in the present, rather than considering their future income, consumption, and savings. Their results hold true for the COVID-19 pandemic, as measured up until the end of 2021.

The testable hypotheses related to the lifecycle hypothesis are:

- Is there a strong positive relationship between income and wealth accumulation?
- Do cross-sectional age-Wealth Index profiles from the survey data resemble the same asymmetrical hump-shape?
- Are there strong positive relationships between patience and/or risk aversion and wealth accumulation, and how do they compare to other determinants of wealth?

Comprehensive wealth: incorporating non-financial assets into the equation

One of the most influential evolutions of the meaning of wealth is the introduction of comprehensive wealth. This concept, developed and popularized towards the end of 20th century, takes the definition of wealth beyond quantifiable, tradable assets, such as net

worth. Notably, it incorporates the idea of human capital, which includes education, physical and mental health, skills, experience, social connections, and other aspects. Arrow et al. (2012) refer to these aspects as 'non-tradable' contributions to wealth: they all contribute to wealth in a more holistic sense, but their monetary value is difficult, or in some cases impossible, to calculate. The idea originally developed at the scale of entire economies (e.g., the education level or the health of a national population), but it holds just as true for individuals. For example, holding a university degree and being in good health clearly allow people to generate net worth (e.g., through being able to work and save), but they also constitute components of comprehensive wealth in their own right.

However, contrary to net worth, which can be objectively measured in a single figure, the complexity of measuring comprehensive wealth poses a statistical challenge. The notion of 'shadow pricing' is used by Arrow et al. (2012) to work through this problem. An illustrative example is the attribution of a value to a certain educational certification by totalling the additional lifetime value of wages it could be expected to generate. These types of calculation are necessarily dependent on assumptions and forecasts, and they vary significantly across distinct groups (e.g., areas of study) or geographies (e.g., a degree might 'generate' more income in Ontario than in Quebec). Crucially, they also measure *potential* wealth in some cases, rather than actual realized wealth. While $1,000 in a chequing account has a face value, this is not the case for the value of a bachelor's degree if the holder has chosen not to work or has chosen to work in a different domain.

As it relates to measuring human capital, Abraham and Mallatt (2022) identify three approaches used by economists. The first two methodologies – the cost approach and the income approach – both rely on the conversion of a human capital asset into a quantifiable financial unit. For example, an individual's investment in their skills could either be measured by the monetary expenses associated with the investment (e.g., the tuition fees) or by the income the investment generates (i.e., the additional earnings from a promotion that can be obtained with the skills). Due to data constraints, these first two approaches will not be used in this thesis, though could offer potential avenues for further research. The third approach – the indicator methodology – relies on using a methodologically consistent data item to measure a given aspect of human capital: a widely used example

would be ‘years of education’. The indicator approach has the advantage of being less impacted by these assumptions, unlike the cost and income approaches, and it facilitates comparisons between many different individuals in the same socioeconomic context. Human capital indicators cannot be directly compared with other measures of capital (notably monetary measures), but they can be constructed straightforwardly and integrated into indexes of comprehensive wealth with an appropriate statistical methodology. Integrating social and human factors into wealth is clearly advantageous if we are interested in measuring the sustainability of wealth. Certain human capital assets have a remarkable feature: their durability. Professional experience cannot be ‘lost’ in the same way that financial wealth can (e.g., a house losing its value during a market downturn).

The testable hypotheses related to comprehensive wealth are:

- Which aspects of human capital are determinant of wealth?
- Are there strong positive relationships between education and/or general ability and wealth accumulation?
- Is there a way of evaluating the role of human capital with respect to other determinants of wealth?

Propensity to plan: linking wealth accumulation to a key cognitive skill

Recent interest in the relationship between the propensity to plan and wealth can be traced to a seminal paper by Ameriks et al. (2003), who argue that an individual’s engagement in financial planning can have a sizeable positive impact on their level of wealth. The authors provide the following theoretical framework: an individual’s level of ‘control skills’ (e.g., the propensity to plan) influences their ability to solve ‘control problems’, issues that are related to the discordance between long-term goals (e.g., wealth accumulation) and present-day actions (e.g., saving). Even if they have the same desire or ambition to accumulate wealth, it is argued that an individual with a strong propensity to plan is more likely to succeed at doing so than an individual who is less inclined to plan. There is an interesting link with the lifecycle hypothesis here. The patience and self-control included into the discount factor are part of ‘control problems’: the “issue’ is that

a dollar spent today means that one dollar less will be available to spend tomorrow. However, propensity to plan is part of 'control skills', which include the ability to devise a plan, make progress on that plan, and take corrective action if the desired goals are not met.

Ameriks et al. (2003)'s findings were remarkable: a one-standard-deviation increase in self-reported propensity to plan led to a 39% in net worth, or a 48% increase in gross financial assets, with other covariates such as education, income, and marital status taken into account. Psychometric factors such as risk aversion and patience were also tested, but had no significant impact on wealth accumulation. Further empirical studies have gone further and deepened our understanding of the propensity to plan and its relationship with wealth accumulation. In one of many studies on retirement planning, Lusardi and Beeler (2006) found that in a sample of 7,515 baby boomers from the Health and Retirement Survey in the US, those who did not engage in financial planning retired with 20% to 45% less wealth than those who did. Lynch et al. (2010) focused on borrowing, finding in a questionnaire sample of 105 adults that a one-point increase in an individual's propensity to plan (on an arbitrary scale) resulted in a 15-point increase in their credit score. More recently, Lee et al. (2019) used a sample of 6,157 respondents from the National Financial Well-Being Survey to show that propensity to plan has a strong relationship with financial well-being and that this effect increases when an individual has greater financial knowledge.

As a determinant of wealth, propensity to plan poses a measurement challenge: simultaneity. Although we can hypothesize that a higher propensity to plan leads to wealth accumulation, it is also the case that wealth accumulation leads to a higher propensity to plan. An illustrative example is an individual who unexpectedly receives a considerable sum of money and seeks out the advice of a financial planner. The relationship is bi-directional, which makes it more difficult to evaluate, because the effect of planning on wealth must be isolated from the impacts going in both directions. While Ameriks et al. (2003) used an instrumental variable approach to take this into account, some empirical studies – such as Lee et al. (2019) – use regression techniques that do not account for the simultaneity. As a result, the results could be biased because the estimators contain

information about both directions of the relationship. The statistical details and solutions are discussed further in Section 5.1.

The methodology used by Ameriks et al. (2003) has been reproduced in other contexts over the past 20 years. Recent examples include Brunetti et al.'s (2022) study, in Italy finding a positive relationship between financial planning and financial knowledge, or Hou and Schuler's (2022) German study that showed a non-linear relationship between financial literacy and wealth (i.e., diminishing returns). This thesis is the first example of a study where the relationship between wealth accumulation and propensity to plan is examined both before and after the onset of the pandemic, and so contributes novel findings to the body of work on the economic impacts of the COVID-19 shock on individuals. Additionally, it uses a uniquely constructed wealth and financial well-being indicator as its dependent variable, which captures a different perspective on wealth than the 'net worth' definition used in other research.

For the purposes of this thesis, planning is taken to be a 'choice' variable (i.e., within individuals' control) and it is taken to mean 'economic planning' (i.e., not just financial planning in the strict sense, but planning of use of resources). Although individuals might have differing innate abilities when it comes to planning, it could be argued that someone who wished to improve their abilities could do so. For example, financial resources and knowledge are much more freely available now than before, as is more practical information about the components of a financial plan, how to set goals, and which financial instruments are appropriate for a given situation. It is important to recognize that many groups have been disadvantaged in this respect, for example due to lack of access to financial knowledge.

The testable hypotheses related to the propensity to plan are:

- Is there empirical evidence from the survey data that economic planning impacts wealth?
- What is the strength of the relationship between economic planning and wealth?
- Does this relationship change after the onset of the pandemic?
- How does planning compare to other determinants of wealth from the literature?

Financial literacy and wealth: widening our view of human capital

Financial literacy, a concept pioneered by Annamaria Lusardi in the early 2000s, is defined as the "ability to process economic information and make informed decisions about financial planning, wealth accumulation, debt and pensions" (Lusardi & Mitchell 2014, 6). As a skill, financial literacy can be understood as a component of human capital (Huston, 2010; Lusardi and Mitchell, 2014; Bucci et al., 2023), and therefore is embedded within comprehensive wealth. Individuals dedicate efforts to increasing their ability to accumulate wealth through, for example, targeted use of financial instruments. The positive relationship between financial literacy and wealth is intuitive and self-reinforcing: individuals with higher levels of knowledge are better positioned to accumulate wealth and in doing so, further improve their financial literacy and experience with different investment vehicles.

The empirical evidence for the relationship between financial literacy and financial well-being is wide-ranging. Letkiewicz et al. (2015) find indeed that wealthier individuals in Canada are more prone to seeking out and paying for financial advice, which contributes to a bi-directional relationship between affluence and financial planning. Boisclair et al. (2017) find a significant relationship between financial literacy and the propensity to plan for retirement, which is most particularly demonstrated in individuals with a stronger understanding of risk diversification (a key aspect of investment). In their adaptation of the lifecycle model to include financial knowledge, Lusardi et al. (2017)'s results suggest that financial literacy plays a key role in explaining wealth inequality (in a net worth sense) in the United States. Gathergood and Disney (2011) find that British high-literacy individuals are less likely to experience unmanageable debt, which has a clear positive impact on wealth. Although financial literacy has received academic attention since the 20th century, the number of empirical studies since 2000 has increased dramatically and is becoming ever more diversified (Santini et al., 2019).

Governments, educational institutions, and actors in the private sector are increasingly paying attention to the role of financial literacy, a sign of its growing importance in the modern understanding of what it means to be wealthy. In Canada, the federal government created a Financial Literacy Month in 2011, which was dedicated to the promotion of

initiatives to support the first of an ongoing series of five-year financial literacy strategies. The theme of the current programme, which runs from 2021-2026, is resilience, defined as "the ability to adapt or persevere through both predictable and unpredictable financial choices, difficulties, and shocks in life" (Government of Canada, 2022). In the context of the individual-level economic shocks associated with the COVID-19 pandemic (unemployment, drastic drops in business income, unexpected expenses etc.), this resilience is of particular relevance, as economic uncertainty and precarity continue to prevail in richer economies. The relationship between financial literacy and resilience is intuitive: individuals with greater knowledge are better prepared to weather economic shocks, for example to lower debt or higher emergency savings. Generally speaking, more financially literate individuals are also much more likely to have developed a financial plan.

The relationship between financial literacy and wealth is widely accepted and has been confirmed by empirical studies from any different countries. However, there is resistance to what has been deemed "financial literacy discourse". One notable example comes from (Pinto & Coulson 2011), two Canadian researchers who argue that by attributing responsibility of financial literacy to individuals, governments can minimize their own responsibility to promote fair distribution of wealth and to create conditions in which individuals can realistically achieve economic well-being. Their view of the Canadian government's financial literacy initiatives is that it sends a message (explicitly or tacitly) that individuals are responsible for their own level of wealth, for example, through engaging in planning or through becoming financially literate. As a result, the policy does not fully recognize the role of 'non-choice' determinants of wealth such as inheritance, sex at birth, and simple luck. Members of groups that suffer systemic discrimination in their access to financial resources – notably people on a low income, women, Aboriginal Canadians – might not be able to realistically improve their situations through the 'choice' of financial literacy and economic planning. To assume that they can do so is to ignore the role of social injustice in economic outcomes.

However valid this critique may be, it is possible, and economically desirable, for individuals to act in a way that increases their own potential for wealth and their actual

wealth, while also holding public actors accountable for good governance and economic policy. As wealth inequality persists at both the individual and community levels, the importance of these responsibilities cannot be overstated.

The testable hypotheses related to financial literacy are:

- Is there empirical evidence from the survey data that financial literacy impacts wealth?
- Are there interactions between financial literacy and planning as it relates to their impact on wealth?

Cognitive and non-cognitive factors: integrating psychology and behaviour into wealth

The label "cognitive and non-cognitive factors" is a broad term, covering psychological, behavioural, and biological factors that could reasonably impact an individual's ability to accumulate wealth. Cognitive skills require intellectual effort and are considered to be 'hard skills' insofar as objective measures exist for them; examples include thinking, reasoning, and recall. The category also includes educational attainment. Non-cognitive skills, referred to as 'soft' skills, are more closely related to personality traits and are more difficult to quantify. These skills are often analysed with the acronym OCEAN, used to denote Openness to experience, Conscientiousness, Extraversion, Agreeableness and Neuroticism, which emerged from the field of psychology (Mosca & McCrory, 2016). Empirical studies have confirmed the theoretical relationships between cognitive and non-cognitive skills and wealth accumulation, many of which are intuitive such as conscientious individuals accessing higher occupational statuses (Exley et al., 2022). There is also overlap between these skills and the other theories, e.g., financial literacy and reasoning; propensity to plan and conscientiousness. By some definitions, the propensity to plan and financial literacy would be considered as cognitive skills.Cognitive and non-cognitive skills are included in this research due to the richness of the dataset, which offers granular insights into several different psychometric areas. Although the true breadth of this body of literature is not covered in this thesis, it is a relevant domain within behavioural economics and has benefitted in the past few years from an increasing number of large-scale empirical studies (Dawson, 2023)

A testable hypothesis related to cognitive and non-cognitive skills is:

- Which psychological and behavioural attributes from the survey data show a relationship with wealth?

4. Measuring Wealth

This section presents the rationale, methodology, and results for the first part of the thesis: the development of a novel indicator to measure wealth. The problems associated with traditional wealth measures, the difference between wealth and income, and the advantages of developing a new indicator will be addressed. The resultant Wealth Index will then serve as a dependent variable for the analyses completed in the second part of the thesis, which is an examination of the role of the propensity to plan in determining wealth.

4.1 What is wealth?

Particularly as it relates to individuals, Brzozowski et al., (2010, 6) say that "wealth data is hard to come by in Canada", and wealth data problems are at the heart of the first part of this thesis. This sub-section introduces the different issues surrounding collection of high-quality data about individual wealth, and the biases that result from measurement error. While it is futile to aspire to a 'perfect' indicator of wealth, the idea of building a novel 'Wealth Index' is to work around some of these issues and to test hypotheses about the determinants of wealth with a different type of measure. Some progress has been made in recent years in this respect: in the Canadian context, the Survey of Financial Security (SFS) asks specific and detailed questions on assets, liabilities, and other aspects of wealth. Since its establishment in 1999, it has been administered five times at irregular intervals, but is becoming increasingly frequent (Statistics Canada, 2020). Given that it has not been administered since the onset of the COVID-19 pandemic, it is unsuitable for a comparison with the results of the data analyzed for this thesis but is further analyzed in the Discussion as a tool for measuring wealth in Canada.

Although it is one of the most abundant terms in economic theory and literature, the term 'wealth' is conceptualized and measured with surprising versatility. Defined in its simplest numerical sense, wealth is the net present value of an economic entity – be that an individual, a family, a business, or an entire economy – calculated by subtracting total liabilities from total assets. This is often referred to as 'net worth' and has a strict financial

definition that can be calculated and demonstrated[3]. As far as meaning is concerned, 'wealth' is synonymous with 'affluence' and 'prosperity', both of which indicate the presence of significant material resources while also implying that such wealth is sustainable.

The concept of durability takes on particular relevance when we consider what it means for an individual to be 'wealthy' compared to having a high income. As a stock variable, wealth is reported at a singular point in time, contrary to income, a flow variable, that requires specification of a time period (most often yearly) in order to be interpretable. Even if we know that an individual has a high income, that is no guarantee of past or future revenue, while wealth tends to have a connotation that the individual has accumulated richness in the past and that it is durable into the future. Wealth and income are frequently conflated, notably in popular parlance where the terms 'rich' and 'poor' do not generally differentiate between high income or high wealth (e.g., a 'rich' person who has inherited wealth might have little or no stream of income, while a 'rich' person who earns a high salary might consume virtually all their revenue and therefore have little wealth). The relationship between income and wealth is clear and is the cornerstone of the neo-classical lifestyle theory, because in that model, an individual's revenue is used to save, thus building wealth. However, as Kennickell and Woodburn (1999) show, the relationship between income and wealth is not necessarily intuitive because it is tempered by several other factors that are often difficult to observe, such as inherited wealth. Indeed, Brzozowski et al. (2010) find that in Canada, the correlation between income and wealth is just 0.35, indicating a moderate relationship at best. This result makes it clear that rather than being synonymous with wealth, income is better viewed as a component or a determinant of wealth.

When asked spontaneously, individuals generally know their annual income, but are much less likely to know their level of wealth at a given moment, which poses a significant measurement problem for microeconomic research on wealth. Several reasons explain

[3] Although there are nuances in exactly how net worth is calculated, such as different approaches to evaluating assets, the key idea is that there are formulaic methodologies that exist to compute a value for net worth, which can be compared between economic entities.

this knowledge gap. Firstly, there is an informational dimension, whereby individuals are not able to precisely recall their debt or asset balances (Juster et al., 1999). This is particularly the case for longer term liabilities and assets such as mortgages and retirement savings, where balance-checking behaviour tends to be less frequent than current accounts (e.g., only with the arrival of quarterly or annual statements), though individuals might also have difficulty recalling their checking account or credit card balances too. Secondly, value uncertainty can influence knowledge of an individual's wealth when the exact value of an asset is not easily quantified: this is notably the case when it comes to real estate. Given structural market changes as well as the specificities of individual properties, homeowners may not be able to estimate a fair market value of their home with precision. The Canadian housing market, one of the world's most dynamic, is a pertinent example of this, particularly in highly populated cities such as Vancouver and Toronto where real estate prices are more volatile and differ substantially from municipal evaluations. Finally, a comprehension problem exists around the subject of wealth: many individuals are not aware of what constitutes wealth (i.e., what is included and excluded in the definition of 'net worth'). Wealth-related concepts such as interest rates and time discounting are poorly understood by the general public (Lusardi et al., 2017; Lusardi & Mitchell, 2017). On the contrary, annual income or revenue is required information on an annual tax return, and most people can recall the figure spontaneously. As a result, it appears that more studies on individual financial well-being focus on income rather than wealth instead[4].

[4] This is a qualitative observation based on the higher frequency of studies using 'income' as a dependent variable or research subject than 'wealth'. As an example, a Google Scholar search for "determinants of income" yields 14,800 results, while "determinants of wealth" yields 1,840.

4.2 Methodology: the Wealth Index

The following methodological sub-Section provides details on the statistical approach used to create the Wealth Index, which is based on the methodology pioneered by Filmer & Pritchett (2001) used by the United States Agency for International Development (USAID) in their construction of socioeconomic indexes around the world. The section also introduces the concept of repeated cross-sectional design for the interpretation between the 2017 and 2021 samples. It concludes with an overview of the key limitations.

Building the Wealth Index

Principal component analysis (PCA) is a research technique with several applications, two of which will be used in this research. The first, dimensionality reduction, is a way of simplifying or condensing a dataset with more than two axes; in this explanation, we consider a space with *k* dimensions (i.e., variables) and *n* observations (i.e., unique responses). Individual observations form a 'cluster' of points that correspond to their values on each of the *k* variables. PCA calculates the mean of all the points in the *k* by *n* space and then recentres this point at the origin, moving the entire cluster of points with it. From this mean point, the principal components (PC) can be established. These are a series of vector projections from the origin: the first principal component, called PC1, is the projection that best fits the data from an error minimization perspective. In other words, it is the vector that captures most of the variance between the observations. Each of the subsequent components begins at the origin and projects outwards: PC2 is the second best fit, PC3 the third, and so on. In contexts like this research, the PCs themselves can take on an interpretable meaning; by identifying the correlation of each variable with the PC, it can be deduced that they can represent a certain characteristic found within the dataset. The dimensionality reduction objective in this case is to identify a PC that represents wealth.

As it relates to the second use, index construction, the output of PCA can be used to create a construction formula for an index containing variables and weightings, as shown in equation (1). Interpretation of the individual PCs is based on two aspects: eigenvalues and factor loading score. The eigenvalue of a PC indicates the extent to which it captures

overall variation in the dataset, because it is the vector of best fit in the cluster of points. Necessarily, PC1 has the highest because it has the lowest mean standard error, and so has the highest eigenvalue. With each additional component (PC2, PC3, PC4, etc.), the eigenvalue declines. Factor loading scores represent the correlation coefficients between each principal component and each of the *k* variables in the dataset. Expressed from -1 to +1, these scores are useful to deduce the relationship of each variable with the component. In this example, the PC of interest will be the one which has a high correlation with variables indicating wealth. Factor loading scores play a second role specific to index construction. Let an index take the form shown in equation (1):

$$I_i = w_1 k_{1i} + w_2 k_{2i} + w_3 k_{3i} + \cdots w_n k_{ni} \quad (1)$$

I represents the index value for a given individual *i*, which is built out of *n* different variables ($k_1, k_2, k_3, \ldots k_n$). In most indexes, the variables are weighted according to their relative importance in determining the outcome of interest. The weightings ($w_1, w_2, w_3, \ldots w_n$) can be obtained from taking the factor loading score of a given variable with respect to the PC of interest. The logic is intuitive: if a PC can be identified as representing wealth, a variable with a high absolute factor loading score will have a stronger relationship with wealth and should be assigned a higher weighting in the calculation of the index. Weightings can be positive or negative.

The use of PCA for constructing socioeconomic wealth indexes is well-documented (Filmer and Pritchett, 2001; Krishnan, 2010; Li et al., 2012). The field of international development has been prominent in this type of research design, primarily due to lack of reliable data or low data quality with respect to wealth and income in developing countries. In these contexts, large numbers of wealth proxy variables are collected in order to build a portrait of household affluence, and then PCA is used to determine the weightings assigned to the indicators used, with greater weightings representing more important factors that explain variation in affluence.

The methodology for this thesis is derived that of Filmer and Pritchett (2001), who used data from a national household survey in India, in order to calculate an Asset Index for

each respondent. Given that their aim was to measure the relationship between school enrolment and wealth level, their study lent itself well to the analysis of multiple factors determining affluence, as is the case in this research. More detailed guides to replicating the methodology have since been updated by the Demographic and Health Surveys (DHS) Program run by USAID, and were studied in order to build the Wealth Index used for this thesis, following Rutstein (2015). That being said, certain contextual differences required adaptation of the approach. In the DHS studies, PCA is used on a series of variables that each represent ownership of a certain asset (e.g., a car, heating, sanitation, etc.), which are all considered to be direct components of household wealth. The factor loading scores of the first PC act as weightings because, by design, they imply a correlation with affluence. In this dataset, there is not an intuitive link between all of the *k* variables and level of wealth. As a result, rather than taking the weightings from the first PC with the highest eigenvalue, an exploratory analysis is performed beforehand in order to identify the PC that best represents wealth. Only then do they factor loading scores have an interpretable meaning that can be used to construct the Wealth Index. Although the variables used in this research are of several different types (binary, ordinal, etc.), the methodology required that they be standardized to have mean 0 and a standard deviation of 1. This avoids certain variables unduly affecting the PCA results simply because of the way in which they are recorded.

Repeated cross-sectional design

Repeated cross-section data is "created where a survey is administered to a new sample of interviewees at successive time points" (UK Data Service, 2015), meaning that this survey data lends itself particularly well to this type of analysis. Although individuals cannot be 'followed' through time as in a longitudinal survey, repeated cross-sectional research designs have the advantage of being logistically more cost-efficient as a way of obtaining data about a large group of people, and do not run the risk of attrition over a long period of time. Blanchard et al. (1977) identify three types of effects that can be deduced from a repeated cross-sectional study. Firstly, age effects are observed as the population grows older, and apply to all cohorts regardless of period. This type of effect takes age as a determinant for a certain outcome, for example, following neoclassical theory, we would expect to see that, for most ages, the average wealth of individuals aged

n+4 in 2021 was higher than that of *n*-year-olds in 2017. The second type of observation is period effects, which are changes that occur at a particular point in time and impact all age groups and cohort effects. The pandemic falls into this category and lends itself well to a repeated cross-sectional design because it has such variable effects on distinct groups, while directly affecting the entire Canadian population. Finally, cohort effects are those that affect populations born at different times and can be understood as the impact of being part of a certain generation. A wealth-related example would be that millennials (born between 1984 and 1995) have a lower median level of net worth than baby boomers (born between 1945 and 1960) had at the same age (Pew Research Center, 2019; Gruijters et al., 2023). The timespan of the dataset precludes an analysis of cohort effects in the context of this thesis, though it remains a relevant avenue for future research on individual wealth.

Certain criteria must be met for a repeated cross-sectional design to be effective. In the case of the data produced by the Lifestyle Survey, this is the case; the criteria and the justifications for considering them fulfilled are listed in Table 2:

Table 2: Dataset criteria for a repeated cross-sectional design

Comparability criterion	Comments
Methodological consistency	The survey methodology is essentially identical between years.
Independent sampling	The sampling is completed independently between years on a representative portion of the Canadian population. There is a negligible probability of some repeat sampling.
Same survey	Although some questions differ, the variables of interest are available in all the survey years.
Year indication	The survey year is clearly identified in the dataset.
Unique identifier	All participants are assigned a unique identifier in each year.

Source: UK Data Service (2015)

Limits of this approach

Regarding the use of PCA for the construction of a Wealth Index, the richness of the dataset is at once an advantage but also a limitation. Contrary to Filmer and Pritchett's (2001) study, where all of the PCA input variables had a clear association with wealth because they all represented asset ownership, the relationships in this dataset are more speculative and diverse in nature. Certain variables, such as income and education, have clear associations with wealth from both the literature and intuition. Others, such as worrying about money, are less well-defined in their relationship with wealth. Subjective decisions were made in order to establish a correlation cut-off of the factor loading scores. As a result, this Wealth Index has a greater degree of subjectivity than those in which a list of assets is inputted, and the resultant wealth index is compiled based on individuals' responses. As it relates to the two samples, the PCA was run on the 2017 sample only, with the construction formula then applied to the 2021 sample. While this allows for a more statistically sound comparison of wealth changes between the two years (because 2017 is established as a 'baseline' and the same measurement technique is used in 2021), the principal components and the factor loading scores yielded different results in 2017 and in 2021. In order to ensure the robustness of the results, a 'shadow analysis' was performed with an alternative version of the Wealth Index, constructed using a different PCA approach. This is discussed in detail at the end of Section 4.3.

As far as repeated cross-sectional design is concerned, its key limitation is that it remains a simulation of a longitudinal study rather than a true analysis of the same individuals over the course of several years. Consequently, the results do not explicitly contain micro-level changes to individuals, but rather speak to macro-level differences across an entire population. In a sample size as large as this dataset, a quasi-longitudinal design does have a decent degree of reliability but must still be interpreted with caution in this context. Furthermore, as a research technique, it is significantly rarer than the other methods used in this research, namely PCA and multivariate regression. Consequently, there is less literature against which to compare the methodology and results of the repeated cross-sectional design results.

4.3 Results

PCA results

Executed on over 400 variables, the PCA yielded definitive results with respect to wealth. Table 3 shows each of the principal components (PCs) 1 to 5 and their eigenvalue, which measures the proportion of total variation that they explain. It also shows commentary on what the likely vector represents within the dataset.

Table 3: Extract of PCA results from the global sample and qualitative analysis

PC	Eigenvalue	Variance captured	Analysis
Comp1	25.64	5.84%	Vector for trend-adopting behaviours and knowledge
Comp2	13.22	3.01%	Inconclusive/unclear
Comp3	9.92	2.26%	Vector for wealth/affluence (stronger)
Comp4	9.13	2.08%	Vector for ambition and self-development
Comp5	6.78	1.65%	Vector for wealth/affluence (weaker)

The PCA was run on the entire sample of 40,602 individuals. Only the first five principal components are shown in this table.

Within these PCs, a diverse amount of information is captured and reduced into a simpler form. As an example, the first component has the highest eigenvalue and therefore accounts for the most variation – 5.8% – in the data. In this case, the variables with the highest absolute correlation (i.e., the maximum positive and minimum negative factor loading scores), clearly identify that PC1 holds information about trend-adopting behaviours. At one end of the spectrum, we observe variables such as 'I like to buy the newest brands/styles', 'I spend a lot on clothes for myself', and multiple variables associated with travel. At the opposite end, we observe 'I am not planning on making any technology-related purchases', 'I do not go to the cinema', and 'I regularly relax at home'. Given the nature of the Lifestyle Survey, which is administrated in a marketing and advertising context, this result is unsurprising. There does not appear to be any information related to wealth contained within this vector. From this, we can presumptively label PC1 as the 'trend-adopting' vector, then repeat the analysis for the remaining PCs.

Two components clearly stood out as relating to wealth: PC3 and PC5, which had eigenvalues of 9.92 and 6.78 respectively and explained a total of 3.9% of the variation in the dataset. While this is low in absolute terms, the combined explanatory power of components 3 and 5 makes affluence the second most important source of variation, after trend-adopting behaviours. In both these cases, several income variables clustered at one end of the component spectrum, notably 'household income', 'personal income', and 'income distribution', showing a clear association with affluence. Given that many different variables repeat themselves between the PCs, the component with the highest eigenvalue, PC3, is retained as being the best representation of wealth. While the full results to the PCA are displayed in Appendix 2, they will not be explored in deeper detail in this section because of the broadness of this level of analysis.

The Wealth Index

The next stage requires the selection of variables and weightings to construct a Wealth Index. In their meta-analysis of PCA research, Vyas and Kumaranayake (2006) argue that there is no clear rule on how many variables should be retained for the index, and that discretion can be used according to the context and the available data.

In the case of this research, certain selection criteria can be ascertained from the thesis objectives and the hypotheses presented in Section 3:

1. Retain variables with a factor loading score higher than a certain threshold;
2. Include variables covering the different dimensions of wealth reviewed in the theoretical frameworks;
3. Avoid variables that are used as covariates in Ameriks et al.'s (2003) paper on the propensity to plan, to ensure that a comparison can be completed.

With these criteria in mind, the index components retained are summarized in Table 4:

Table 4: Factor loading scores for the retained variables

Variable	Factor Loading Score (on wealth principal component)
Household income distribution	0.2032
Own house/apartment	0.1483
Self-assessed financial situation	0.1120
'I worry a lot about money'	-0.1112

The formula for the construction of the Wealth Index is therefore as follows:

$$WI_{i,t} = 0.2income_{i,t} + 0.15ownhome_{i,t} + 0.11financialsituation_{i,t} - 0.11worryaboutmoney_{i,t} \quad (2)$$

Figure 1 and Table 5 show the distribution and summary statistics for the Wealth Index as built by equation 2, on an arbitrary scale of 0 to 4.3.

Figure 1: Distribution of the Wealth Index variable

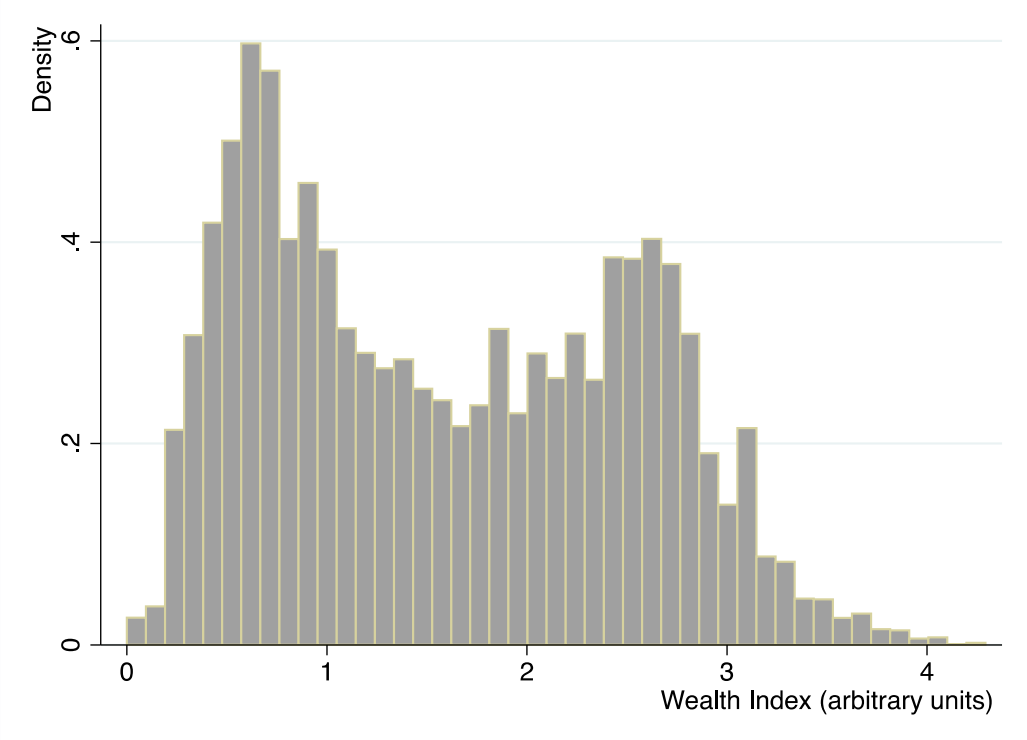


Table 5: Summary statistics for the Wealth Index and its component variables

VARIABLES	N	Mean	St. Dev.	Min.	Max.	Q25	Median	Q75
Wealth Index	36,473	1.598	0.913	0	4.292	0.762	1.500	2.424
Household income	36,473	5.17	3.949	1	16	1	5	9
Homeowner dummy	40,602	0.578	0.494	0	1	0	1	1
Financial situation	40,602	3.496	1.061	1	5	3	4	4
Worry about money	40,602	3.341	1.178	1	5	3	3	4

Like any constructed index, the Wealth Index takes arbitrary values based on the components and their weightings, and as such, it is a relative scale from 0 to 4.3. Although the construction formula output yielded both positive and negative values, the score was adjusted linearly by adding the absolute minimum to each value to ensure that all values were positive. The avoidance of values above and below zero aids interpretation and facilitates certain statistical approaches, such as the use of logarithms. The Wealth Index has an asymmetrical bimodal distribution, with two local maxima: the lower is at approximately 0.8 and the higher is at approximately 2.7, meaning that two-thirds of the overall sample falls between these two values. Of the remaining third of the population, 23% are 'low wealth', falling below the first maximum, while 10% are high wealth, falling above the second maximum.

Use of aggregate data to verify the results of the Wealth Index

As shown in Table 6, we observe an increase in wealth of 2.8% over the four-year period, which is far below the Statistics Canada estimates for net worth increase in Canada, but more representative of Canadians' self-reported level of financial well-being (see Section 6).

Table 6: Evolution of the components of the Wealth Index between 2017 and 2021

Component	2017 mean	2021 mean	% change
Wealth Index	*1.58*	*1.62*	*+2.8%*
Household income	5.05	5.31	+5.2%
Homeownership dummy	0.617	0.531	-13.9%
Self-assessed financial situation	3.47	3.53	+1.7%
Worry about money	3.32	3.36	+1.2%

Results shown for the full sample of 22,274 respondents (2017 sample) and 18,328 respondents (2021 sample).

The main driver of this growth was household income, which rose by 5.2%, and is the highest-weighted component of the group. The income variable is reported on a distribution (1 = under $25,000 to 15 = $500,000 or more) and is in line with we could expect in terms of growth over four years and a major economic shock. Self-assessed financial situation, which grew by 1.7% from 3.47 to 3.53, also contributed to the growth of the Wealth Index. This change moves the mean closer to 'comfortable' (a score of 4 on the Likert scale) than 'coping' (a score of 3). The negative effects acting on wealth are homeownership and worrying about money. The percentage of respondents owning their homes declines from 61.7% to 53.1% between 2017 and 2021. This is a significant brake on Wealth Index growth, given that it restricts access to the most popular wealth-generating asset. Financial anxiety grew slightly from 3.32 to 3.36, which is surprisingly low given the severity of the economic impact of the pandemic. In Canada, this could be explained by the provision of emergency measures by both the federal and provincial governments, which contributed to keeping individuals' and businesses' cashflows stable amid the lockdown and closure of certain segments of the economy.

Empirical evidence from the Wealth Index related to theories of wealth

Figure 2 shows the age-Wealth Index profiles: the mean Wealth Index value for each age, taken over the entire sample (i.e., respondents in both 2017 and 2021). We can clearly observe a steep increase in wealth between the ages of 15 and 26, when mean wealth increases from 0.67 to 1.5 in just ten years. After this inflection, the Wealth Index rises more steadily to a reach peak of 1.9 between the age of 39 and 43. From 44 years old and upwards, we observe a gradual decrease in the Wealth Index, albeit with some

fluctuations, and then a slightly steeper drop-off starting around retirement age (62-65 years old).

Figure 2: Age-Wealth Index profile for the global sample (N = 36,473)

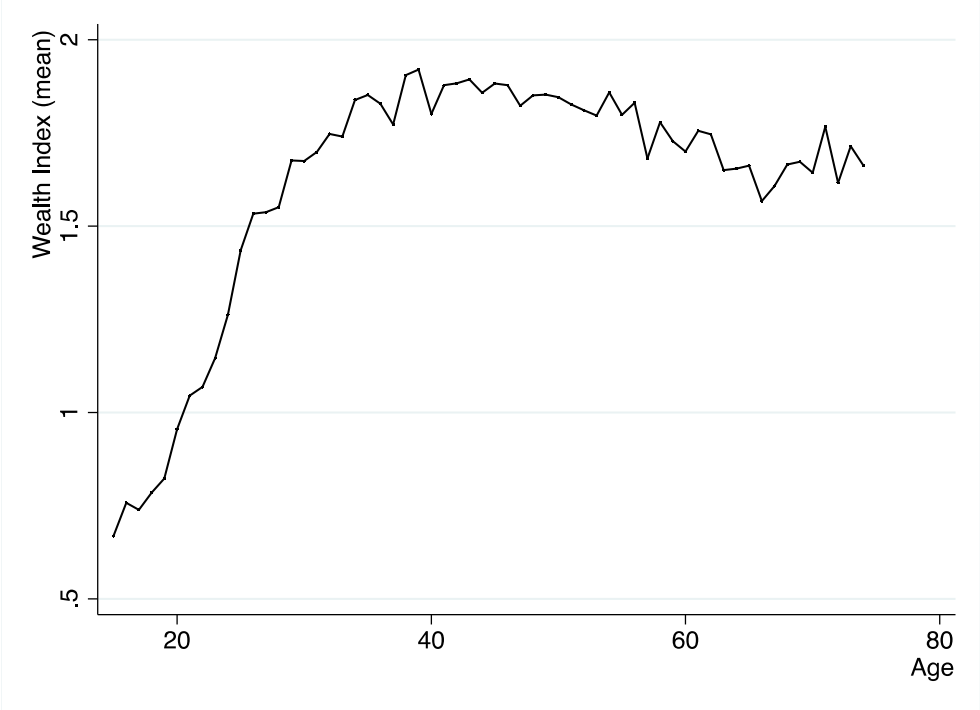


Figure 3: Age-Wealth Index profiles for the global sample, by gender (N = 36,473)

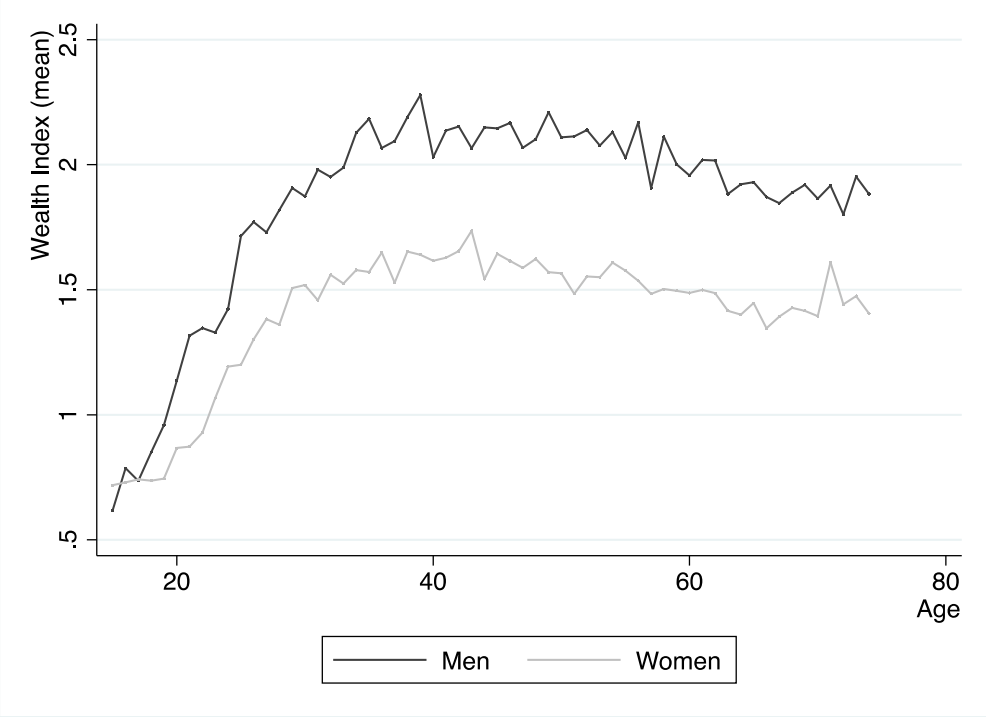


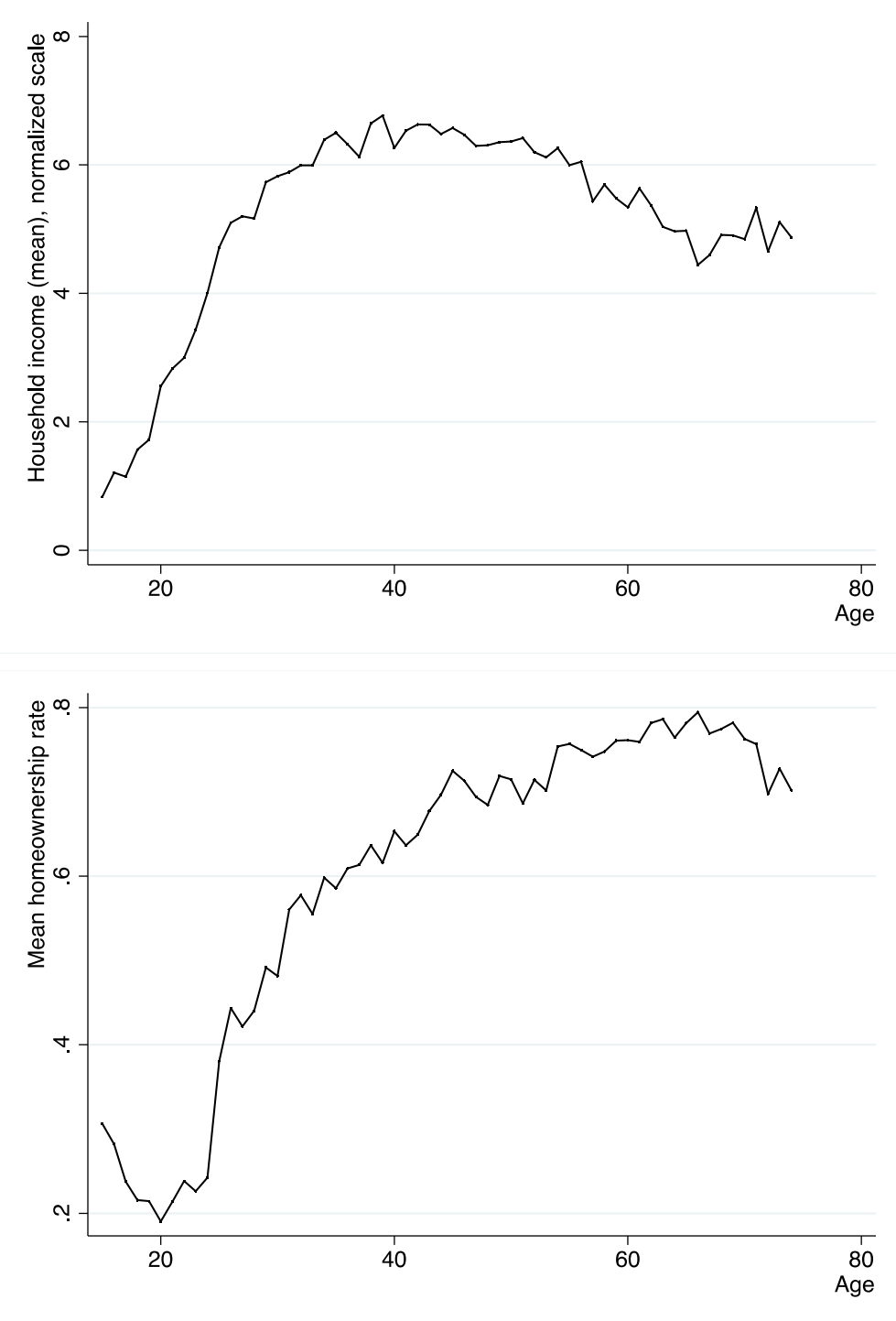
Figure 3 provides the gender breakdown of the age-Wealth Index profile for the global sample. The same broad pattern is discernable for men and women: a steep growth in wealth, followed by a more gradual increase, and then a very gradual downwards-leaning plateau. At every age, men have a higher Wealth Index than women, though this gap is very narrow at a young age and then increases throughout individuals' twenties, reaching a plateau from the thirties onwards.

One key explanation for wealth from neoclassical theory is that wealth accumulation occurs primarily through income and age, whereby revenue earned is either used to consume or to save, and wealth is accumulated through savings. As an individual moves through their lifecycle, they tend to both earn more and have more accumulated in savings, which results in a wealth peak just before retirement and a decline in the later years of life. With respect to accumulation, the lifecycle model predicts that wealth accumulation accelerates over a lifetime: An individual with a lower income will dedicate a much greater proportion of their revenue to consumption, but over time, this ratio shifts in favour of saving, and already-accumulated wealth grows through interest-related gains.

Firstly, the 'hump' described in the lifecycle hypothesis is present in the age-Wealth Index profiles: we can observe an increase, followed by a peak, followed by a gradual decline. These support the predictions of the lifecycle hypothesis. However, the positions of the inflection points in the Wealth Index profiles are different from what the lifecycle hypothesis would suggest. Rather than an accelerating rate of wealth accumulation that peaks just before retirement, we observe a decelerating rate of wealth accumulation and a middle-age peak. While the lifecycle hypothesis suggests that the majority of wealth is accumulated in the later years of life, when an individual contributes more of their revenue to saving, the evidence from the Wealth Index suggests most wealth is developed between 15 and 26, and then the rate of increase declines into the middle age. To understand this pattern, we can look for explanations in age profiles for each of the components of the Wealth Index: household income, home ownership, and the three subjective measures of financial well-being. As with all age profiles, a limitation of the Wealth Index is that does not simulate one individual lifetime, because the age profiles are aggregated across a population at one moment in time rather than following a cohort of individuals over the

course of their life. However, the aggregate portrait is of analytical value and shows the patterns we would expect from the lifecycle hypothesis.

Figure 4: Age-profile for each Wealth Index component



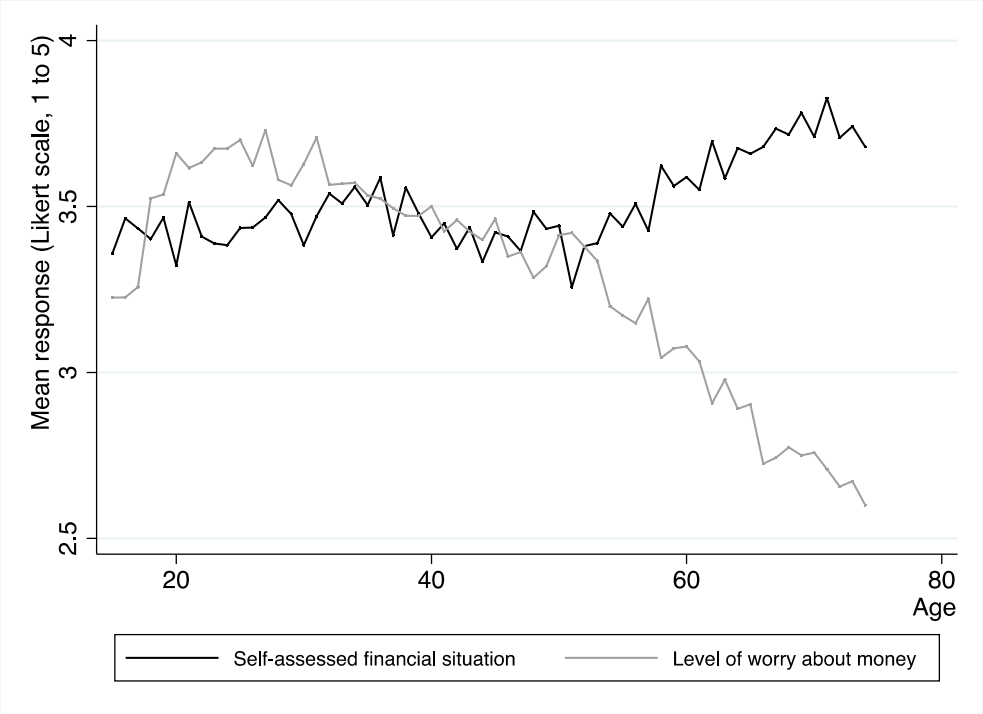
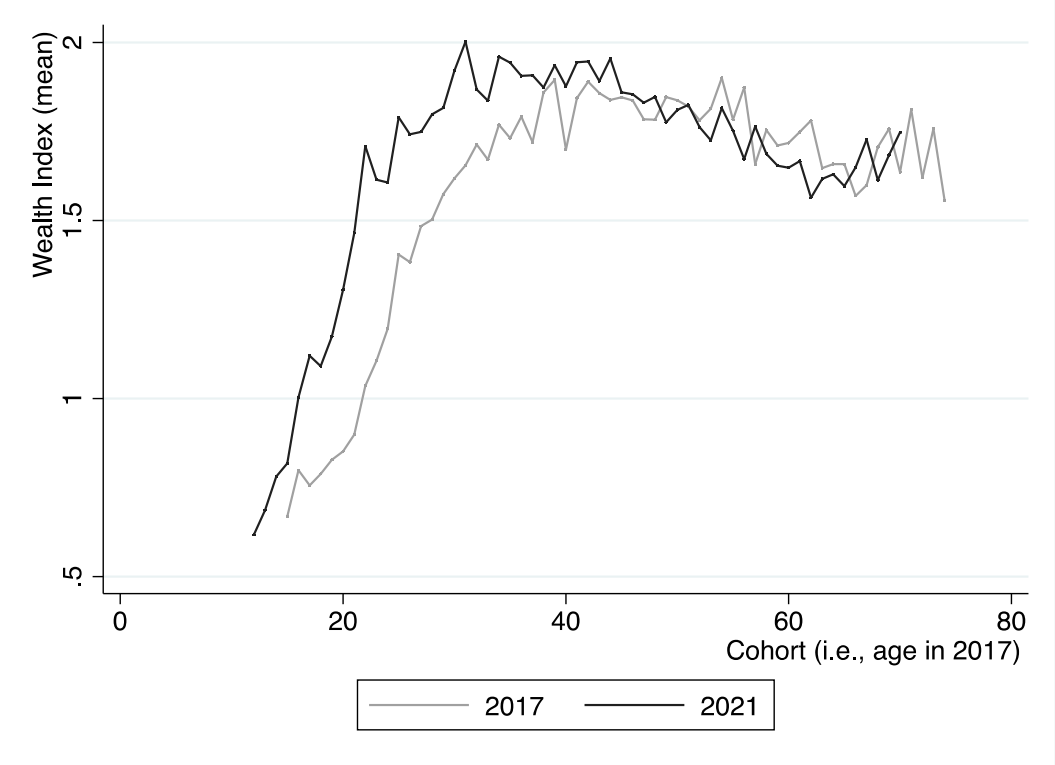


Figure 4 shows the age profiles for the four components of the Wealth Index: household income (the highest weighting), homeownership, self-assessed financial situation, and worry about money (the lowest weighting). As highest-weighted component, it is logical to see that the age-income profile best matches the age-Wealth Index profile. For homeownership, we see a clear reflection of the pattern observed in the age-Wealth Index profiles, but the shape is different. Following the steep rise during an individual's twenties, homeownership rate continues to steadily increase, even into retirement. The initial dip appears to be a response error, whereby individuals under 20 reported 'owning' their homes if their parents were homeowners. The impact on the Wealth Index results is negligible, so no adjustment is necessary. The two subjective financial well-being measures show the most fluctuation. Individuals experience the most financial anxiety in their twenties, which decreases notably during the middle age despite no change in self-assessed financial situation. The indicators both suggest the presence of an inflection point at around 55 years old where individuals develop financial confidence: their self-assessed financial situation improves, and they worry less and less about money. This result is reassuring for the internal validity of the survey.

The gaps between the predictions of the lifecycle hypothesis and the Wealth Index results can be further illustrated using the repeated cross-sectional (quasi-longitudinal) design, whereby we can compare the wealth of population groups aged *x* in 2017 with that of the population aged *x*+4 in 2021, in order to ascertain the age effect on wealth. Figure 5 shows the mean Wealth Index, by cohort, where cohort is taken to represent a respondent's age in 2017. In this sense, a respondent aged 25 in the 2017 sample is in the same cohort as a respondent aged 29 in the 2021 sample. Age effects are represented by the vertical Euclidean distance between the two lines, for a given cohort. Age effects are significantly greater between 18 and 30 than they are from 34 onwards, at which point they begin to stabilize (i.e., the impact of one year on wealth for a 45-year-old is much lower than it is for a 25-year-old).

Figure 5: Cohort-Wealth Index profiles, 2017 sample and 2021 sample



One clear theoretical comparison can be made with these results. While lifecycle theory emphasizes material wealth, the Wealth Index takes income, homeownership, self-assessment, and financial anxiety into account. In this sense, wealth accumulation occurs much earlier on, as individuals see higher income increases, first-home home buying, and

a (modest) improvement in their own assessment of their finances. In this sense, the Wealth Index could be interpreted as a Wealth *Potential* Index, which would represent an individual's accumulation of *ability* to generate further wealth, rather than the wealth itself. For example: someone who buys property earlier in life will be better equipped to generate more wealth with that property than someone whose first purchase comes later. The idea of a Wealth Potential Index is explored further in Section 6.

Figure 6: Age-Wealth Index profiles, by highest level of education reached

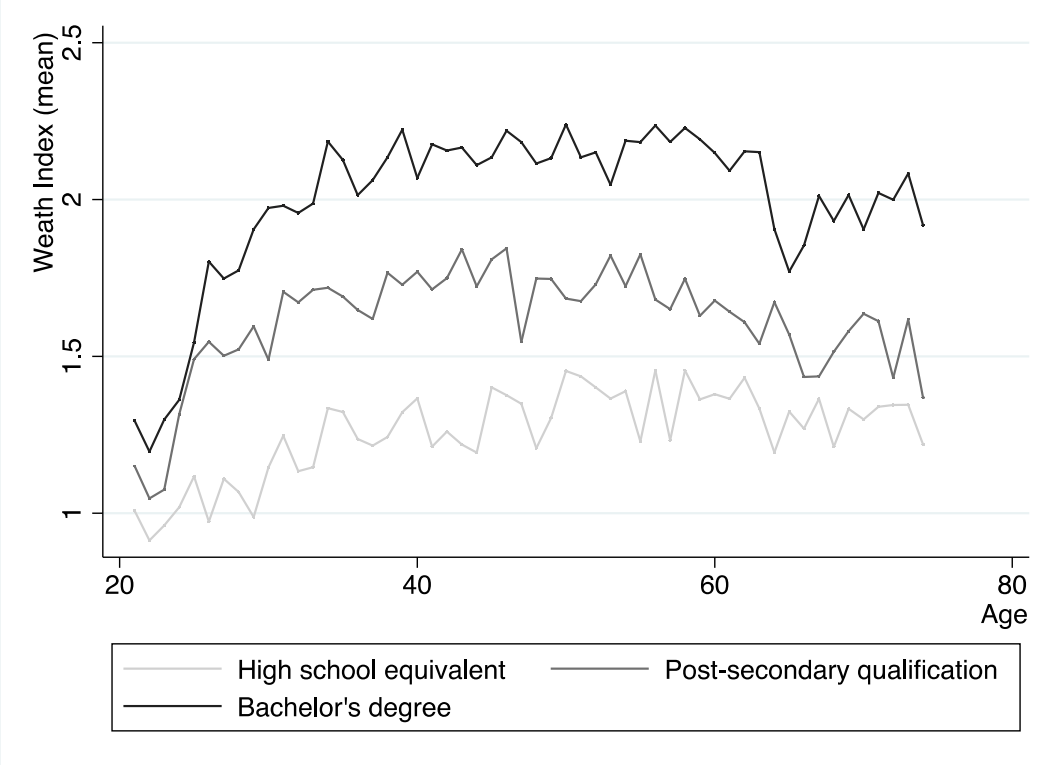


Figure 6 shows the age-Wealth Index profiles for the global sample, divided by level of education reached. As far as human capital is concerned, the age-Wealth Index profiles are coherent with the notion of comprehensive wealth when they are broken down by level of education. Across the lifecycle, individuals with more years of education enjoy consistently higher levels of wealth, with the most significant gap between high school diploma holders and individuals with a post-secondary qualification. As would be expected, the gap is narrow at a younger age, and increases as individuals reap the benefits of education in the labour market.

Robustness of the Wealth Index and 'shadow analysis'

The quality and the limitations of the Wealth Index are explored in more detail in Section 6. One key feature of PCA index construction is the arbitrariness of many of the steps: which variables to include in the original PCA set, how many to retain from the select component(s), the cut-off point for the factor loading scores, and the statistical manipulation of the values computed by the construction formula. The innumerable different permutations of the Wealth Index give rise to an important question: what if the index were constructed a different way? To respond to this limitation, the results associated with a different formation of the Wealth Index are included in Appendix 3 as a 'shadow analysis': each of the results from the main body of the thesis are repeated, in some case with appropriate methodological adjustments. The decision of which Wealth Index to choose was largely based on the comparability with other studies, notably Ameriks et al.'s (2003) propensity to plan results. Given that the alternative Wealth Index value included variables that Ameriks et al. (2003) use as covariates (e.g., education), a fair comparison could not be made between the results. The shadow analysis is intended to highlight the reliability of the results of the main index while also demonstrating different approaches that can be used to develop alternative measures of wealth. The results of the subsequent analyses are broadly consistent, indicating the validity of the statistical approaches used.

Table 7 highlights some of the key differences between the two wealth indexes. The Alternative Wealth Index uses the full sample from both years, which is more representative of the population but makes the index less appropriate for a comparison between 2017 and 2021, because it 'blends' the conditions between the two years. In the main Wealth Index, only one round of PCA is used to identify the PC that represents wealth. In the shadow analysis, a second round of PCA is executed, taking the 90 variables with the highest factor loading scores from the first round and running a PCA with the restricted set. The Alternative Wealth Index also has a lower factor loading score cut-off, with all variables having a correlation of +0.075 and more or -0.075 and less included in the construction formula. While this appears to be a low correlation in absolute terms, it is high in relative terms due to the diversity of the dataset. The core Wealth Index has a slightly higher cut-off, and consequentially, has fewer variables in the construction

formula. A final key difference was the decision to include income in the index. The logic of including household income in the core Wealth Index is that it is a key determinant of affluence and should be considered in assessment of an individual's wealth. The logic of excluding it in the Alternative Wealth Index is to favour a more comprehensive, non-financial view of wealth and leaving space for variables such as education and health.

Table 7: Selected differences between the core and alternative indexes

Decision	Wealth Index	Alternative Wealth Index
Sample used for PCA	2017 only	Full sample (2017 and 2021)
Number of rounds	1 round only	2 rounds
Factor loading score cut-off	|0.10|	|0.075|
Variables included	4	9
Income included	Yes	No

As illustrated in the results sections and in Appendix 3, the indexes perform very similarly and indicate the same broad patterns about wealth and about planning. As a robustness check, this is very reassuring when it comes to drawing conclusions and practical implications from the data. It is particularly powerful because the indexes are very different in their composition (only two variables – home ownership and financial anxiety – appear in both) and because several aspects of the PCA process are different. In their meta-analysis of PCA studies, Vyas and Kumaranayake (2006) argue that many of the specific research design decisions are arbitrary and tend to be based on data availability. The similarity in the results of the two analyses gives support to this idea.

5. Economic Planning and Wealth Accumulation

This section will cover the second part of this thesis, concerning the following research question: how does planning influence individuals' wealth accumulation? Various instrumental multivariate regression techniques will be used in order to estimate the causal relationship between the propensity to plan and wealth, generally using the Wealth Index constructed in the previous section as a dependent variable. The first part of this section will outline the methodology and the second part will present the results and economic interpretations. The question of the pandemic shock will be addressed by comparing the 2017 and 2021 samples and identifying how the shock may have influenced the relationship between planning and wealth.

5.1 Methodology: the wealth-planning relationship

Overview

Let equation 3 represent the model that represents the research question at hand:

$$w_{i,t} = \alpha_0 + \alpha_1 \mathbf{x}_{i,t} + \alpha_2 \boldsymbol{p}_{i,t} + \varepsilon_{wi,t} \quad (3)$$

The variable $w_{i,t}$ is a measure of wealth for individual *i* at time *t*, the vector $\mathbf{x}_{i,t}$ is a set of socio-demographic controls known to determine wealth, while $\boldsymbol{p}_{i,t}$ represents the propensity for economic planning. An error term ε_w is specific to this model in that it captures idiosyncratic variation in wealth. The dependent variable *w* will be measured directly with the Wealth Index, which has a minimum of 0 and a maximum of 4.3. As for the vector **x**, given that the dataset is rich in socio-demographic information, it will be based on the covariates used by Ameriks et al. (2003), including age and sex. Given the highly urbanized nature of Canada and the concentration of wealth in its four largest cities (notably through homeownership), a dummy is added for individuals living in rural areas. Additionally, Ameriks et al. (2003) did not use any data on ethnicity in their original study, either by choice or due to data availability. In this thesis, a dummy is also added for respondents who confirmed their ethnicity as white, testing findings from studies that non-white Canadians are more likely to experience worse financial outcomes when it

comes to income derived from wealth (Block et al., 2019), and even when they have higher levels of educational attainment (Statistics Canada, 2023).

The choice of variable to represent $\boldsymbol{p}$ is critical. As it relates to defining planning, examples from the literature inform this approach. Ameriks et al. (2003) asked respondents a specific question about their engagement in financial planning, i.e., do you have a financial plan or work with a financial adviser. Lusardi and Mitchell (2008) formulated questions about retirement planning, such as "Have you developed a plan for retirement saving?" and "How often have you been able to stick to this plan?". In this case, the survey data available contain two variables that indicate economic planning ability, which corresponds to an individual's propensity to plan. The 'planning variables' are identified as P1 and P2:

P1: "I need to be sure I've considered all the options before I make a purchase decision"

P2: "I budget very carefully when doing the household shopping"

Within the survey data, these variables are the most appropriate to act as the dependent variable. Given that two variables are available, one will be used in the regression model, and the other will be reserved to assess the robustness of our results. As both planning variables are measured on the same 5-point Likert scale, they are mathematically comparable with each other.

OLS correlation analysis

In the first instance, an initial set of regression results is obtained by running an ordinary least squares (OLS) model on equation (3). This coefficient of the planning variable in this estimation will represent a 'naïve' indicator of the impact of planning on the Wealth Index value. Two reasons explain why the result will not likely be representative of the true causal impact of planning on wealth. Firstly, there is an issue of simultaneity: the relationship between economic planning and wealth is likely bidirectional, in that planning influences wealth but wealth also influences planning. Indeed, Ameriks et al. (2003) use this logic, arguing that wealthier individuals are much more likely to engage in financial planning because of their ownership of more complex assets. In this example,

the opposite could even be true, whereby lower wealth individuals tend to show a higher propensity to plan economically because they are limited in their resources, compared with higher wealth individuals who spend more spontaneously because they have the flexibility to do so. This is particularly the case with P2, where a careful budget could indicate a low level of affluence. Secondly, omitted variables could also lead to inaccurate estimates, whereby a key determinant of wealth is excluded (inadvertently) from the model specification and is embedded in the error term, causing a bias. In this example, people with high wealth might see economic planning as unnecessary because they consider themselves to be comfortable already. People with lower levels of wealth might plan more in an attempt to improve their financial situations. These correlations, which are not causal links, could have an impact on the estimator in the model and the result impacts the strategy required to evaluate the causal effect of planning on wealth. In these cases, an omitted variable could be anything excluded from the model that has a direct impact on wealth accumulation.

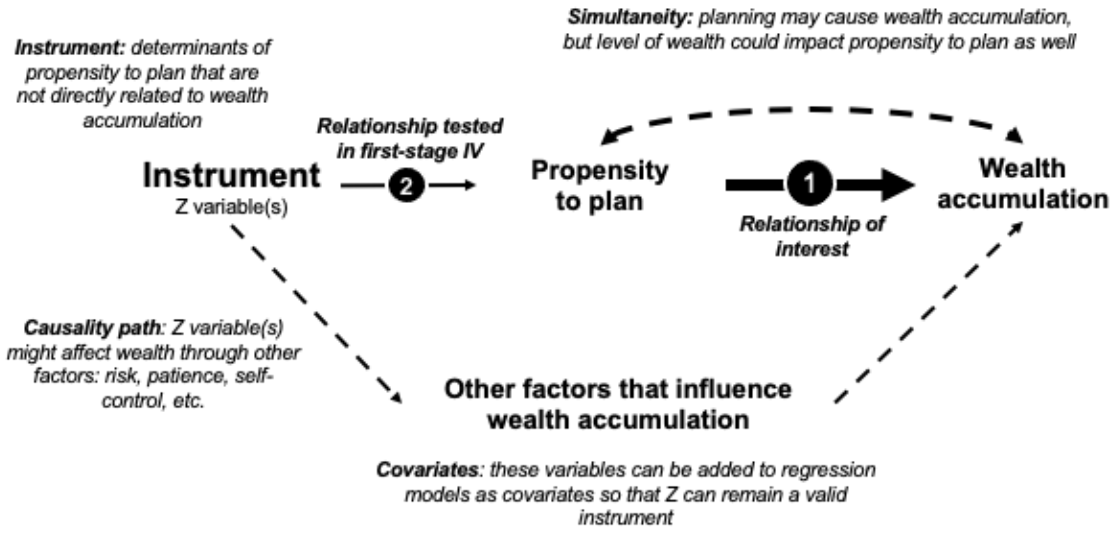
The OLS regression is run in order to create a 'baseline' impact of planning on wealth, net of the impacts of simultaneity or the omitted variable bias, which could have either a positive or negative on the coefficient of the planning variable. The coefficient generated by OLS will then be compared to the results of the IV regression; the greater the difference between the two results, the stronger the implied bias is. A negative bias (i.e., the OLS coefficient underestimates the impact because it is lower than the IV coefficient) likely indicates omitted variable bias or negative simultaneity, where planning and wealth impact both each other but in opposite ways. A positive bias would suggest that there is positive simultaneity in which the influence of planning on wealth and wealth on planning is the same direction (i.e., both positive or both negative).

IV regression design

One solution to the problem of simultaneity is instrumental variable regression. As illustrated in Figure 7, when a model features reverse causality, an instrument (z) that influences the independent variable exclusively through its effect on the dependent variable of interest can be used to isolate and evaluate the true causal relationship,

removing the estimator bias of a model with reverse causality. The OLS coefficient will be used as a baseline against which to compare the results obtained by IV regression.

Figure 7: IV regression schema



Finding an instrumental variable to measure the relationship shown as arrow 1 is a statistical challenge, and when selecting an instrument, two key conditions must be fulfilled. Firstly, the instrument must be relevant, defined as having a significant relationship with the independent variable of interest (i.e., planning). In the schema, this ensures that arrow 2 makes sense. The fulfilment of the first condition is evaluated through an OLS regression of the model in equation (4) below:

$$p_{i,t} = \beta + \beta_1 \mathbf{x}_{i,t} + \beta_2 \mathbf{z}_{i,t} + \varepsilon_{pi,t} \quad (4)$$

The term $\mathbf{z}$ indicates the instrumental variable candidate(s) that are being evaluated. In order for the relevance condition to be met, a significant, non-zero coefficient for β_2 must be observed, indicating that the instrument has a statistical relationship with the planning variable. Additionally, the result of an F-test of a regression model with only the selected instrument can suggest whether there is a weak instrument problem, whereby the z-variables are insufficiently related to planning and the IV regression results are not meaningful. Stock et al. (2002, p. 522) provide a rule of thumb F-statistic threshold when

testing multiple potential instruments, which will be considered in the interpretation of the results.

The second condition is exogeneity: the instrument must be uncorrelated with the error term of the original model such that $Cov(z, \varepsilon_w) = 0$, where z is the instrument and ε_w is the error term as it relates to wealth. This is the notion that the only mechanism though which the instrument impacts wealth is through its relationship with planning (proved as part of the first condition), which is required to satisfy the exclusion restriction. There can be no direct causality between the instrument and economic planning. While there is no formal test for this condition, it is possible to hypothesize about potential sources of correlation between the instrument candidate variables and the error term ε_w. Caution is necessary at this stage, because the result of this test relies implicitly on estimates that could be biased if the instrument is not valid.

Evaluation of the relevance and the exclusion restrictions is performed with the full set of controls in mind. Furthermore, a robustness test is conducted on the final results by running the regression for planning only, then with the first covariate, then with the first and second covariates, and so on, in order to examine the results as each incremental independent variable is added. This approach ensures that the relationships identified in the results and the logic of the instrumental variable design hold true across different model specifications.

Instrument selection

From the behavioural variables available in the dataset, three candidates were retained for the instrument, based on their ability to fulfil the conditions described in the methodology.

The variables are all on the 5-point Likert scale:

$\boldsymbol{z_1}$: "I like to follow a well-organized routine"

$\boldsymbol{z_2}$: "I make plans for my friends and for myself"

$\boldsymbol{z_3}$: the inverse of "I like to enjoy life without needing a plan"

Contrary to planning and formal education, which are cognitive, these variables all represent non-cognitive abilities: a propensity towards a routine can be seen as the quality

of being organized and disciplined; making plans for oneself and friends speaks to organizational skills and collaboration; needing a plan can be viewed as being structured. While the two types of abilities are clearly intertwined, the use of instruments from the non-cognitive family of skills is beneficial in that it is more likely to avoid unwanted interactions with the planning variable and the other covariates. This point will be revisited in the robustness checks of the results.

The appropriateness of the instruments is tested by equation (5):

$$p_{i,t} = \alpha + \beta_0 \mathbf{x}_{i,t} + \beta_1 \mathbf{z_1}_{i,t} + \beta_2 \mathbf{z_2}_{i,t} + \beta_3 \mathbf{z_3}_{i,t} + \varepsilon_{pi,t} \quad (5)$$

To ensure consistency, each of the variables represents responses to an agree/disagree statement on the Likert scale, in this case from 1 (strongly disagree) to 5 (strongly agree) and a neutral option 3 in the middle (neither agree nor disagree). These variables all have the potential to affect to an individual's propensity to plan, condition 1 (relevance), but have the critical characteristic of not being directly associated with wealth accumulation, condition 2 (exclusion restriction).

5.2 Results

In this sub-section, the results of the IV strategy will be presented in three parts. Firstly, a baseline planning-wealth relationship is established through simple OLS regression. Then, the candidates for the instrumental variable (z_1, z_2, and z_3) are evaluated in the first-stage IV results. The retained instrument is then used in the estimation of the full IV model, which gives the second-stage results.

OLS correlation analysis results

Table 8 shows the results and the explanatory notes give details about the covariates that are used in further models in this section. The coefficient of P1 is -0.0348 and -0.0821 for P2, suggesting that a one-point increase in planning would *decrease* wealth by 0.03 or 0.08 units. In other words, moving from ‘agree’ to ‘strongly agree’ on the planning question would have a negative correlation with wealth. Keeping in mind that the scale for planning goes from 1 to 5, this is an economically insignificant effect.

Table 8: Baseline OLS results for the effect of planning on wealth

VARIABLES	(1) Planning as P1	(2) Planning as P2
Planning	-0.0348***	-0.0821***
	(0.00407)	(0.00375)
Single dummy	-0.234***	-0.233***
	(0.00866)	(0.00861)
Education level	0.162***	0.160***
	(0.00225)	(0.00224)
Age	0.0579***	0.0594***
	(0.00156)	(0.00155)
Square of age ÷ 100	-0.0572***	-0.0586***
	(0.0176)	(0.0175)
Male dummy	0.409***	0.402***
	(0.00805)	(0.00801)
White dummy	0.112***	0.108***
	(0.0104)	(0.0103)
Rural dummy	-0.160***	-0.158***
	(0.00845)	(0.00840)
Constant	-0.328***	-0.187***
	(0.0358)	(0.0347)
Observations	35,803	35,803
R-squared	0.317	0.325

Standard errors in parentheses. Significance of coefficients is shown as follows: *** p<0.01, ** p<0.05, * p<0.1. Results for the entire sample, less respondents who did not provide an answer. The planning variable is either P1 (carefully considering options before a purchase decision) or P2 (budgeting carefully when doing the household shopping). The dummies for are all binary variables that take 1 if the respondent is single, male, white, or rural, and 0 if not. Education is on a sliding scale of 1 (high school) to 7 (Master's or PhD completed), broadly corresponding to the notion of 'level' of education. Age is in years. See Appendix 7 for a full list of definitions.

IV results: first-stage

Table 9 shows the results for the first-stage OLS regression, with the instrument candidates are in grey. The results with each instrument candidate appearing individually can be found in Appendix 5: Additional Regression Results.

Table 9: First-stage IV regression results

	(1)	(2)
VARIABLES	P1	P2
Single dummy	0.0194*	0.0362***
	(0.0105)	(0.0112)
Education	0.0165***	-0.0172***
	(0.00273)	(0.00291)
Age	0.00510***	0.0216***
	(0.00188)	(0.00201)
Square of age ÷100	-0.0286	-0.0189***
	(0.00211)	(0.00226)
Male dummy	-0.0280***	-0.0886***
	(0.00976)	(0.0104)
White dummy	-0.0587***	-0.0648***
	(0.0125)	(0.0134)
Rural dummy	0.00247	0.0318***
	(0.0102)	(0.0109)
Routine (z_1)	0.157***	0.193***
	(0.00491)	(0.00524)
Plan friends (z_2)	0.0689***	0.0739***
	(0.00465)	(0.00496)
Inverse of 'No plan' (z_3)	-0.0450***	-0.0793***
	(0.00476)	(0.00508)
Constant	2.851***	2.423***
	(0.0470)	(0.0502)
Observations	39,784	39,784
R-squared	0.040	0.054

Standard errors in parentheses. Significance of coefficients is shown as follows: *** p<0.01, ** p<0.05, * p<0.1. Results for the entire sample, less respondents who did not provide an answer. The planning variable is either P1 (carefully considering options before a purchase decision) or P2 (budgeting carefully when doing the household shopping). The dummies for are all binary variables that take 1 if the respondent is single, male, white, or rural, and 0 if not. Education is on a sliding scale of 1 (high school) to 7 (Master's or PhD completed), broadly corresponding to the notion of 'level' of education. Age is in years. See Appendix 7 for a full list of definitions. Each of the instrument candidates is on a Likert scale of 1 to 5. The F-test score for the instrument candidates is provided in the following paragraph.

The results indicate that z_1 – the proclivity to follow a routine – appears to have a stronger relationship with planning than the other two candidates. Surprisingly, the z_3 instrument candidate has a negative coefficient, suggesting that people who do not enjoy life without needing a plan have a higher propensity to plan financially. Although the reason for this relationship is not apparent, it could be the case that individuals who engage in planning have more financial freedom to be spontaneous, or it could be the result of a social desirability bias where the wording of the question – which includes "enjoys life" - encourages people to respond more positively than is actually the case, therefore decreasing the instrument which is the inverse of the response. As for z_2, planning for friends and self, there is a considerable positive relationship with the planning variable and a robust t-value. However, the results from the PCA indicated that wealth accumulation is linked to extraversion, confidence, and a healthy social life. With this in mind, we can hypothesize that z_2 is more likely than z_1 to have a non-zero covariance with the error term of equation (3). The results of an OLS regression with the planning variables on the three instruments give an F-statistic of 511.72 and 635.69 for P1 and P2 respectively. Stock et al.'s (2002) study suggests that a three-instrument first stage OLS model should have an F-statistic of 9.08, indicating that a weak instrument problem is not an issue.

Of the three instruments, z_3 is a less appealing choice, because the effect itself is very small and the sign of its coefficient is the opposite to what was expecting. While z_2 remains an interesting choice, it is a less convincing candidate to fulfil the exclusion restriction condition. The routine instrument z_1 appears to be the best candidate, given its strong positive relationship with planning and intuitively better position to not impact wealth accumulation other than through planning. It is retained to instrument $\boldsymbol{p}$ in equation (3), and the planning for friends instrument (z_2) will be used independently to perform robustness checks. It is important to remember that the 'perfect' instrument does not exist. While there is no obvious link, adherence to a routine could have a direct impact on wealth and therefore a correlation with the error term in the model. Additional rigour and robustness checks are required in the results analysis to ensure that the findings truly support the conclusion that the propensity to plan has an impact on wealth.

IV results: second-stage

Table 10 shows the results of the estimation of equation (4) for the entire sample, both 2017 and 2021. The coefficients of both planning variables P1 (considering options before making a purchase) and P2 (carefully budgeting when doing the household shopping) become positive, take values of 0.152 and 0.125, and are significant at the 1% level. The Sargan test for overidentification indicated that there is no evidence to support the hypothesis that the model is overidentified.

Table 10: Second-stage IV regression results (full sample)

VARIABLES	(1) P1 model	(2) P2 model
Planning (P)	0.152***	0.125***
	(0.0250)	(0.0208)
Single dummy	-0.234***	-0.236***
	(0.00891)	(0.00897)
Education	0.158***	0.162***
	(0.00237)	(0.00234)
Age	0.0568***	0.0552***
	(0.00161)	(0.00167)
Square of age ÷100	-0.0562***	-0.0546***
	(0.00182)	(0.00187)
Male dummy	0.415***	0.423***
	(0.00831)	(0.00858)
White dummy	0.125***	0.125***
	(0.0108)	(0.0109)
Rural dummy	-0.157***	-0.161***
	(0.00870)	(0.00876)
Constant	-0.991***	-0.856***
	(0.0949)	(0.0752)
Observations	35,803	35,803
R-squared	0.277	0.267

Standard errors in parentheses. Significance of coefficients is shown as follows: *** p<0.01, ** p<0.05, * p<0.1. The Sargan statistic was 0.000. Results for the entire sample, less respondents who did not provide an answer. The planning variable is either P1 (carefully considering options before a purchase decision) or P2 (budgeting carefully when doing the household shopping). In both cases, the planning variable has been instrumented for z_1 ("I like to follow a well-organized routine"). The dummies for are all binary variables that take 1 if the respondent is single, male, white, or rural, and 0 if not. Education is on a sliding scale of 1 (high school) to 7 (Master's or PhD completed), broadly corresponding to the notion of 'level' of education. Age is in years. See Appendix 7 for a full list of definitions.

Of the two planning variables, P1 shows a stronger relationship with the Wealth Index and has a marginally higher R-squared score. Given the mean Wealth Index value of 1.6, the P1 result implies that a one-unit increase (e.g., 2 to 3, or 3 to 4) in the propensity to plan increases wealth by 9.5%, all other things being equal. The magnitude of the relationship is striking: when compared with the OLS result of negative 0.0348, the coefficient supports the hypothesis that planning positively impacts wealth and it shows a much greater degree of causality. The changing of the sign from negative to positive

indicates that either negative simultaneity or omitted variable bias (or both) caused the OLS correlation analysis to return a negative result. For example, low wealth could be associated with high economic planning because individuals are more conscious of their spending and engage more in budgeting behaviour. Additionally, the original OLS model specification could be impacted by an omitted variable that does not pose the same problem in the IV model. Importantly, the use of the IV strategy results in the wealth-planning model providing estimates that are much more aligned with the empirical estimates presented in Section 3.

The relative similarity of the coefficients of all the controls between the OLS and the IV models gives additional confidence as to the relationship between the covariates and wealth: the controls show intuitive, well-proven economic patterns: education is a key determinant of wealth (human capital), and wealth increases with age but with diminishing returns (lifecycle theory). Demographic variables follow expected trends: people who are single have lower levels of wealth than those who are married, and there is a wealth gap between men and women of about 6%. The new findings of this research, compared with Ameriks et al.'s (2003) results, are on the wealth gap between white and non-white respondents and urban and rural respondents, both of which showed highly significant coefficients.

Robustness checks

As far as the first-stage results are concerned, a useful robustness check is to test for interactions between the selected instrument, the planning variable, and the other covariates. Table 11 shows the results of OLS regressions with planning and adherence to a routine as dependent variables and the other covariates as regressors.

Table 11: Robustness checks for planning and routine variable interactions

VARIABLES	(1) Planning	(2) Routine
Male dummy	-0.0325***	-0.0587***
	(0.00993)	(0.0100)
Single dummy	0.00158	-0.0585***
	(0.0107)	(0.0108)
Education	0.0243***	0.0344***
	(0.00276)	(0.00279)
Rural dummy	-0.0143	-0.0315***
	(0.0104)	(0.0105)
White dummy	-0.0663***	-0.00914
	(0.0127)	(0.0129)
Age	0.00513***	0.0115***
	(0.00191)	(0.00193)
Square of age ÷100	-0.00436e**	-0.000171***
	(0.00215)	(0.00217)
Constant	3.543***	3.421***
	(0.0400)	(0.0405)
Observations	39,784	39,784
R-squared	0.004	0.012

Standard errors in parentheses. Significance of coefficients is shown as follows: *** p<0.01, ** p<0.05, * p<0.1. Results for the entire sample, less respondents who did not provide an answer. The planning variable is P1 (carefully considering options before a purchase decision). The dummies for are all binary variables that take 1 if the respondent is single, male, white, or rural, and 0 if not. Education is on a sliding scale of 1 (high school) to 7 (Master's or PhD completed), broadly corresponding to the notion of 'level' of education. Age is in years. See Appendix 7 for a full list of definitions.

The OLS results indicate that none of the variables has a very strong effect on planning or routine. Although many of the coefficients are significant at the 1% level, their absolute value is small, with a maximum of just 0.0663 (the white dummy in the planning model). Given that the mean values for planning and routine are 3.699 and 3.625 respectively, this impact is minimal. The extremely low R-squared values, 0.004 and 0.012 compared with 0.277 for the IV model, suggest that the model has little explanatory power despite the inclusion of several covariates.

While the IV results are encouraging and support the hypothesis of the planning-wealth relationship, the initial OLS regression showed the possibility of an omitted variable bias and returned a negative coefficient for planning. As a robustness check on the IV results, additional regression models are run to ensure that the coefficient obtained is not spurious and can also be identified across different specifications. Table 12 shows the results of this check, starting with a simple regression of the instrumented planning variable on wealth, and then adding on each of the covariates one by one.

Table 12: Robustness check of the wealth-planning relationship

VARIABLES	1	2A	2B	2C	2D	2E	3
P	0.314***	0.340***	0.297***	0.144***	0.136***	0.143***	0.152***
	(0.0303)	(0.0296)	(0.0286)	(0.0257)	(0.0257)	(0.0257)	(0.0250)
Male dummy		0.478***	0.469***	0.434***	0.426***	0.429***	0.415***
		(0.00999)	(0.00959)	(0.00849)	(0.00848)	(0.00847)	(0.00831)
Single dummy			-0.453***	-0.352***	-0.361***	-0.347***	-0.234***
			(0.00971)	(0.00866)	(0.00866)	(0.00868)	(0.00891)
Education				0.181***	0.176***	0.180***	0.158***
				(0.00236)	(0.00238)	(0.00237)	(0.00237)
Rural dummy					-0.104***	-0.141***	-0.157***
					(0.00863)	(0.00886)	(0.00870)
White dummy						0.185***	0.125***
						(0.0108)	(0.0108)
Age							0.0568***
							(0.00161)
Square of age ÷100							-0.0562***
							(0.00182)
Constant	0.437***	0.123	0.472***	0.251***	0.354***	0.177*	-0.991***
	(0.112)	(0.110)	(0.107)	(0.0935)	(0.0940)	(0.0957)	(0.0949)
Observations	36,473	36,331	36,014	35,803	35,803	35,803	35,803
R-squared	-0.124	-0.077	0.016	0.235	0.241	0.245	0.277

Models 2A-2E add each of the covariates, one by one, in order of size in the full model. Model 3 shows the full IV model. All models have the Wealth Index as the dependent variable. Model 1 only contains the planning variable as a regressor.

Standard errors in parentheses. Significance of coefficients is shown as follows: *** p<0.01, ** p<0.05, * p<0.1. Results for the entire sample, less respondents who did not provide an answer. The planning variable is P1 (carefully considering options before a purchase decision). The dummies for are all binary variables that take 1 if the respondent is single, male, white, or rural, and 0 if not. Education is on a sliding scale of 1 (high school) to 7 (Master's or PhD completed), broadly corresponding to the notion of 'level' of education. Age is in years. See Appendix 7 for a full list of definitions.

The robustness check results show a consistency that supports the conclusions of the IV models. The coefficient of the planning variable remains positive and significant in each iteration of the model, oscillating between a maximum of 0.340 and a minimum of 0.136 as the other covariates are added.

As the robustness check in Table 11 showed, there were no unexpectedly strong relationships between the propensity to plan or the routine instrument and the other covariates. However, looking more closely at age, although the coefficient of the planning variable is small, we can observe a pattern. Teenagers and young adults have a low propensity to plan, though this increases gradually until it stabilizes at a score of about 3.7 (out of 5) over adulthood. Between 65 and 70 years old, the score declines, likely corresponding to a decreased need or propensity for economic planning during retirement.

Figure 8: Mean propensity to plan, by age

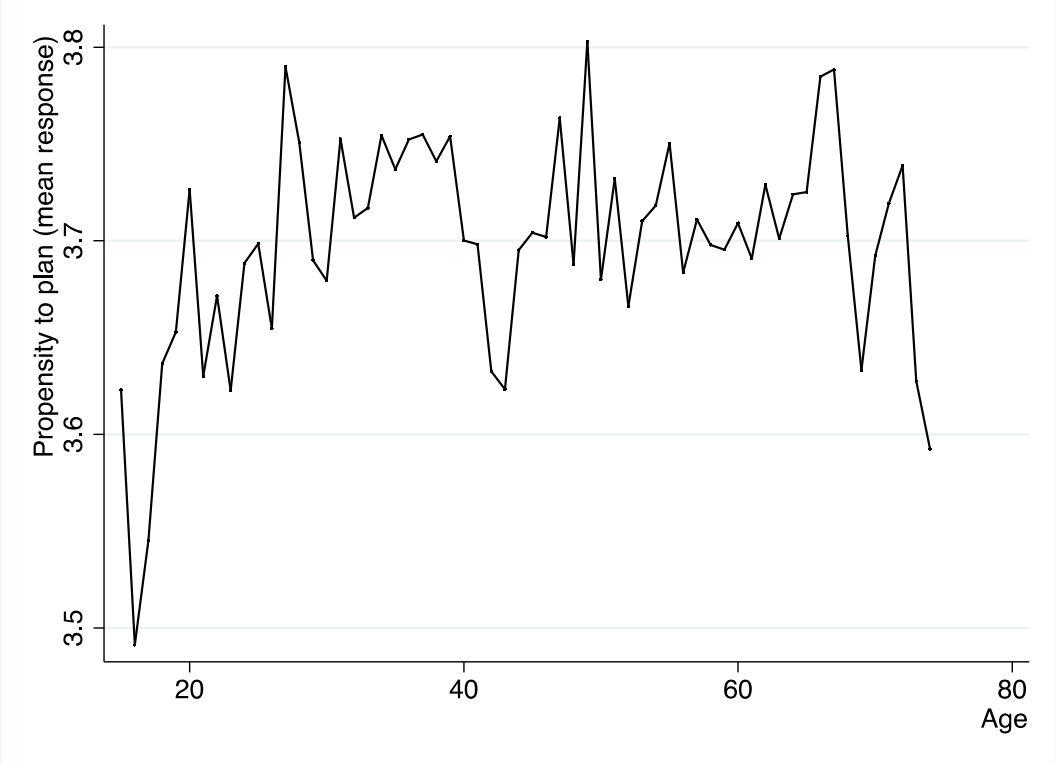


Figure 8 shows the mean value of P for the different age groups in the global sample. While the relationship is highly volatile over time, fluctuating significant from one age to another, we can discern a pattern of individuals increasing their propensity to plan over

time, notably between the ages of 16 and 35, at which point the increase levels off. Planning becomes 'locked in' at this age, without much of a change in trend after that. This point will be examined further in Section 6, with respect to the implications that it might have for improving financial outcomes.

Determinants of wealth: how does planning fare?

The richness of the dataset allows for additional regressions to be run in order to compare the relative impacts of the propensity to plan against other cognitive and non-cognitive factors on wealth accumulation. Table 13 shows the results for three additional IV models, specified by adding one additional variable in order to compare the coefficient to that of planning in the baseline model.

Table 13: Regression results with additional cognitive and non-cognitive factors

VARIABLES	(1) Baseline	(2) Risk	(3) Ability	(4) Impatience
Planning (P1)	0.152***	0.149***	0.113***	0.169***
	(0.0250)	(0.0249)	(0.0260)	(0.0275)
Single dummy	-0.234***	-0.230***	-0.230***	-0.239***
	(0.00891)	(0.00887)	(0.00879)	(0.00898)
Education	0.158***	0.156***	0.154***	0.158***
	(0.00237)	(0.00236)	(0.00233)	(0.00238)
Age	0.0568***	0.0572***	0.0575***	0.0575***
	(0.00161)	(0.00161)	(0.00159)	(0.00162)
Square of age ÷100	-0.0562***	-0.0558***	-0.0563***	-0.0571***
	(0.00182)	(0.00181)	(0.00179)	(0.00182)
Male dummy	0.415***	0.398***	0.399***	0.414***
	(0.00831)	(0.00834)	(0.00831)	(0.00833)
White dummy	0.125***	0.133***	0.124***	0.127***
	(0.0108)	(0.0108)	(0.0107)	(0.0109)
Rural	-0.157***	-0.150***	-0.154***	-0.158***
	(0.00870)	(0.00867)	(0.00858)	(0.00873)
'Other people view me as a risk taker'		0.0609***		
		(0.00370)		
'I have more ability than most people'			0.0668***	
			(0.00493)	
'If I really want something I'd rather buy it on credit than wait'				-0.0317***
				(0.00565)
Constant	-0.991***	-1.172***	-1.070***	-0.954***
	(0.0949)	(0.0940)	(0.0914)	(0.0908)
Observations	35,803	35,803	35,803	35,803
R-squared	0.277	0.284	0.297	0.271

Standard errors in parentheses. Significance of coefficients is shown as follows: *** p<0.01, ** p<0.05, * p<0.1. Results for the entire sample, less respondents who did not provide an answer. The planning variable is P1 (carefully considering options before a purchase decision). The dummies for are all binary variables that take 1 if the respondent is single, male, white, or rural, and 0 if not. Education is on a sliding scale of 1 (high school) to 7 (Master's or PhD completed), broadly corresponding to the notion of 'level' of education. Age is in years. See Appendix 7 for a full list of definitions. Each of the cognitive and non-cognitive factors is measured on a Likert scale of 1 to 5.

Three additional self-reported personality variables were added to the model: propensity to take risks, leadership, and general ability. As with planning, the variables are all reported from 1 to 5 on the Likert scale. In each case, as was expected, there is a significant and positive correlation with wealth, with minimal changes to the coefficients of the other covariates. The coefficients for the planning variable change with each model, which is

likely due to interactions between planning and the additional cognitive and non-cognitive factors added. In each case, the absolute value of the coefficient for the additional factor is lower than the coefficient for planning: 0.149 vs. 0.0609 for risk, 0.113 vs. 0.0668 for ability and 0.169 vs. 0.0317 for patience. In these results, the propensity to plan is a stronger predictor of wealth outcomes than several key behavioural attributes reviewed in the literature. The changing coefficient for planning indicates that our causal IV estimates are marginally affected by the exclusion of these covariates, though these are likely to be correlate with both planning and wealth accumulation.

Instrumental quantile regression

Contrary to OLS or IV regression, which identify the conditional mean of the dependent variable given the covariates, quantile regression (QR) identifies the quantiles along the conditional distribution of the dependent variable. In this context, it will allow us to understand if the relationship between the propensity to plan and wealth holds across individuals of different levels of wealth (rather than just the mean), as well as how it changes at different quantiles of the Wealth Index. To facilitate comparability, the QR is specified in the same way as the IV regression.

Table 14: QR results at each decile of the Wealth Index distribution

Decile	Wealth Index value at decile[5]	Coefficient of P	Coefficient / Wealth Index[6]
10%	0.484	0.127***	+26.2%
20%	0.667	0.109***	+16.3%
30%	0.873	0.134***	+15.3%
40%	1.13	0.143***	+12.7%
50% (median)	1.50	0.128***	+8.5%
60%	1.89	0.125***	+6.6%
70%	2.24	0.148***	+6.6%
80%	2.54	0.152***	+6.0%
90%	2.85	0.141	n/a

Significance of coefficients is shown as follows: *** p<0.01, ** p<0.05, * p<0.1. Results for the entire sample, less respondents who did not provide an answer. The planning variable is P1 (carefully considering options before a purchase decision).

Table 14 shows an extract of the results of the QR; the detailed results for each quantile can be found in Appendix 4. The first column shows the value of the Wealth Index at the decile and the second shows the coefficient of planning on wealth in the QR result at the decile, i.e., the result of interest of the quantile regression for that specific decile. The third and final column is the coefficient divided by the Wealth Index value. With respect to the coefficient values, a clear insight emerges about the robustness of the results of the IV regression: planning impacts wealth across the entire wealth distribution. All but one of the coefficients are positive and significant to the 1% level. This suggests that the conclusions from the IV results are not being driven only by certain individuals with specific levels of wealth, but is in fact a generalized relationship. A specific pattern is difficult to discern, though the general rule appears to be that the coefficient of P is greater

[5] The Wealth Index value cited is the unconditional value at the decile, without controls taken into account.
[6] Unlike in the OLS and IV regressions, this value does not provide a true estimate of the percentage impact on wealth of a one-point impact on planning. Further explanations are provided in the paragraphs following Table 14.

at higher deciles. That said, the coefficient ceases to be significant at the 90% decile, meaning that the statistical confidence in the planning-wealth relationship is lower than an acceptable threshold.

The interpretation of these QR results must be done with caution. The result of interest is not just the coefficient of P, but also the implied percentage impact on wealth of a one-point increase in planning at a given quantile of wealth. To calculate this quantile treatment effect precisely, the potential wealth distribution (i.e., the predicted Wealth Index values) for each quantile would be required. However, given that planning is relatively uniformly distributed across the entire wealth spectrum, and that there are no strong interactions between the planning variable and the covariates (see Robustness checks), we can approximate by dividing the coefficient of P by the unconditional Wealth Index value at the decile (i.e., the observed Wealth Index distribution). It is important to note that this is an approximation of the implied percentage impact on wealth and not a precise calculation as in the previous OLS and IV regression examples (D. Kaplan, personal communication, 23 November 2023)

From these approximations, a very clear pattern emerges with respect to how the strength of the relationship varies with wealth: the lower the Wealth Index value, the higher the relative impact of planning on wealth. To illustrate this insight, consider the 10% and 50% (median) deciles, whose coefficients for P are 0.127 and 0.128 respectively. The implication of this result is that a one-point increase in planning will have virtually the same absolute effect on wealth for an individual in the lowest decile as it would for an individual who has a median level of wealth. However, in percentage terms, the first individual has a Wealth Index value of approximately 0.48, while the second's is about 1.50; the relative impacts of the one-point increase would therefore be +26.2% and +8.5% respectively. In other words, the QR results suggest that lower-wealth individuals have more to gain in relative terms from an increase in their propensity to plan when compared with their higher-wealth counterparts.

It must be kept in mind that the percentages in Table 14 are not formal calculations of quantile treatment effects. However, given the clarity of the results, which show a uniform

pattern of a stronger relationship in lower wealth deciles, combined with the even distribution of planning across wealth levels, it is reasonable to draw the conclusion that an increase in planning is more impactful on wealth for individuals at the lower end of the Wealth Index. Furthermore, an alternative specification is constructed in which the log of the Wealth Index is used as the dependent variable. Taking advantage of the mathematical properties of logarithms, this enables a different approximate calculation of percentage change by multiplying the coefficient of P by 100. The alternative specification provides support for the idea that the relationship is stronger at lower levels of wealth, showing the same pattern as the QR results with the regular Wealth Index. Although the percentage impacts themselves are smaller, the trend is identical: the strength of the relationship increases as the decile decreases. The implied percentage change at each decile is shown in Table 15, and the full results are in Appendix 4.

Table 15: QR results and implied percentage impacts using log of wealth

Decile	Coefficient of P (on log of Wealth Index)	Implied approximation of percentage impact
10%	0.171***	+17.1%
20%	0.134***	+13.4%
30%	0.130***	+13.0%
40%	0.102***	+10.2%
50% (median)	0.096***	+9.6%
60%	0.078***	+7.8%
70%	0.083***	+8.3%
80%	0.076***	+7.6%
90%	0.061**	+6.1%

Implied percentage impact calculated by using the interpretation that a one-point increase in planning causes an increase in wealth calculated by 100 × coefficient of P. This is an approximate result only.

Standard errors in parentheses. Significance of coefficients is shown as follows: *** p<0.01, ** p<0.05, * p<0.1. Results for the entire sample, less respondents who did not provide an answer. The planning variable is P1 (carefully considering options before a purchase decision).

5.3 The propensity to plan and the pandemic shock

Comparative analysis of the IV results

While the dataset does not allow us to follow individuals over years (as a longitudinal design would), it does allow us to compare patterns in 2017 and 2021, between which years the shock of the COVID-19 pandemic occurred. The results in Table 16 show the same regression model run on the 2017 pre-pandemic onset sample and the 2021 post-pandemic onset sample.

Table 16: IV regression results for 2017 and 2021 samples

	(1)	(2)
VARIABLES	2017 sample	2021 sample
Planning	0.145***	0.156***
	(0.0348)	(0.0359)
Single dummy	-0.208***	-0.270***
	(0.0119)	(0.0134)
Education	0.164***	0.151***
	(0.00310)	(0.00366)
Age	0.0514***	0.0631***
	(0.00211)	(0.00250)
Square of age ÷100	-0.0493***	-0.0641***
	(0.00239)	(0.00279)
Male dummy	0.387***	0.437***
	(0.0109)	(0.0128)
White dummy	0.104***	0.151***
	(0.0149)	(0.0158)
Rural	-0.139***	-0.176***
	(0.0113)	(0.0135)
Constant	-0.917***	-1.062***
	(0.134)	(0.134)
Observations	19,420	16,383
R-squared	0.293	0.265

Standard errors in parentheses. Significance of coefficients is shown as follows: *** p<0.01, ** p<0.05, * p<0.1. Results for the entire sample, less respondents who did not provide an answer. The planning variable is P1 (carefully considering options before a purchase decision). The dummies for are all binary variables that take 1 if the respondent is single, male, white, or rural, and 0 if not. Education is on a sliding scale of 1 (high school) to 7 (Master's or PhD completed), broadly corresponding to the notion of 'level' of education. Age is in years. See Appendix 7 for a full list of definitions.

We can observe the increase of the coefficient of the planning variable from 0.145 in 2017 to 0.156 in 2021. Given the mean values of the Wealth Index (1.58 in 2017 and 1.62 in 2021) The implied causal impact rises from 9.1% to 9.6%, which represents an increase of 5.5% in the strength of the relationship. These results are significant at the 1% level in the two subsamples, and the difference between the p-values is relatively small.

Robustness checks

As a robustness check, an alternative specification of the same model was run using a year dummy interaction term. The result shown in Table 17 show that the coefficient of P for the whole sample is 0.143, while the coefficient of the planning and 2021 interaction term is 0.0127, both significant at the 1% level. This gives an almost identical result to the two-sample comparison (the combined coefficient of P and the interaction term is 0.156) and coherent with the conclusion that the impact of planning on wealth became significantly stronger following the start of the pandemic. Given the two sets of results, we can conclude with more confidence that the link between planning and wealth does indeed become stronger after the onset of the pandemic.

Table 17: IV model, alternative specification with year dummy

VARIABLES	
Planning	0.143***
	(0.0249)
Planning × 2021 dummy	0.0127***
	(0.00230)
Single dummy	-0.236***
	(0.00890)
Education	0.158***
	(0.00237)
Age	0.0569***
	(0.00161)
Square of age ÷100	-0.0563***
	(0.00181)
Male dummy	0.412***
	(0.00830)
White dummy	0.129***
	(0.0108)
Rural dummy	-0.156***
	(0.00869)
Constant	-0.984***
	(0.0947)
Observations	35,803
R-squared	0.279

Standard errors in parentheses. Significance of coefficients is shown as follows: *** p<0.01, ** p<0.05, * p<0.1. Results for the entire sample, less respondents who did not provide an answer. The planning variable is P1 (carefully considering options before a purchase decision). The dummies for are all binary variables that take 1 if the respondent is from the 2021 cohort, single, male, white, or rural, and 0 if not. Education is on a sliding scale of 1 (high school) to 7 (Master's or PhD completed), broadly corresponding to the notion of 'level' of education. Age is in years. See Appendix 7 for a full list of definitions.

To further confirm the robustness of the results, the same model was run using the alternative planning instrument (z_2, the proclivity to plan activities for friends). This candidate was not retained because it did not perform as well as the routine instrument, but it was retained for robustness checks. With z_2 the global P coefficient was even higher, taking a value of 0.537. This result suggests that the impact of a one-point increase in

planning on wealth is 33.6%. Furthermore, when the sample was split into the pre- and post-pandemic onset groups, the coefficient of the planning variable was 0.442 in 2017 and 0.634 in 2021, which adds support to the conclusion that the planning-wealth relationship became stronger following the pandemic. It is important to interpret the results of the second instrument with prudence. Having a large social network is related to high wealth, which could mean that z_2 is correlated with the error term in the original wealth model. This would mean that it is a less effective instrument, as it contains additional information about wealth that is not found in the routine variable. The full results of the robustness checks can be found in Appendix 5.

Evolution of the Wealth Index and its components between 2017 and 2021

Another interesting result is obtained by comparing the evolution of the subsamples' wealth between 2017 and 2021, as shown in Table 18.

Table 18: Evolution of mean and median Wealth Index between 2017 and 2021, by propensity to plan

Propensity to plan (P1 variable)	Wealth Index, 2017 sample	Wealth Index, 2021 sample	% change
Mean wealth			
5 (highest)	1.496	1.560	+4.3%
4	1.605	1.655	+3.1%
3	1.587	1.633	+2.9%
2	1.619	1.614	-0.3%
1 (lowest)	1.531	1.500	-2.0%
All respondents	*1.578*	*1.622*	*+2.8%*
Median wealth			
5 (highest)	1.352	1.352	0.0%
4	1.520	1.591	+4.7%
3	1.500	1.520	+1.3%
2	1.554	1.463	-5.9%
1 (lowest)	1.408	1.224	-13.0%
All respondents	*1.500*	*1.500*	*0.0%*

Results for the entire sample, less respondents who did not provide an answer. The planning variable is P1 (carefully considering options before a purchase decision).

Mean wealth levels in the 'highest planners' sample increased by 4.3%, while the 'lowest planners' see a wealth decrease of 2.0%, which shows a significant growth gap of 6.3% between the two ends of the planning spectrum. The overall change in the level of wealth between 2017 and 2021 was an increase of 2.8%. We see a similar pattern when median wealth is used instead, with even more drastic drops in wealth for 'low planners'. These results extend our understanding of the relationship between planning and wealth: there is not only a positive causal impact of a propensity to plan on wealth accumulation, but there is also a connection between planning and wealth change over time before and after the shock.

While the IV regression speaks to the generalized impact of planning and wealth, the comparative results could be interpreted as the impact of planning on economic resilience. With the logic of a repeated cross-section simulation, if we were to track individuals between 2017 and 2021, these results suggest that those who engaged in more planning were better prepared to weather the financial shock of the pandemic, whatever their level

of wealth. In other words, the period effects on wealth over the four years would be more positive for 'high planners', and less positive for 'low planners'.

The results do not appear to be driven by one single component of the Wealth Index. Appendix 6 shows the sample comparisons between 2017 and 2021 for each of the four variables, as well as the percentage change, for planning 'groups' 1 to 5. In terms of household income, high planners are the only group to have a higher income (+1.8%) after the onset of the pandemic than before it. The change in homeownership rates vary significantly between the groups: while it declined for all the population, the decrease was much more pronounced in low planners (-21.2%) than in higher planners (-13.2%) or neutrals (-12.2%). Of the two Likert indicators, change in self-reported financial situation was relatively uniform between the groups, though low planners did experience a greater increase (+1.9%) than higher planners (+1.6%). Conversely, growth in financial anxiety was more pronounced in lower planners (+2.0%) than in higher planners (+1.4%).

These results also speak to a key advantage of using a multi-year dataset: we can compare individuals' propensity to plan with the evolution of wealth. Although the dataset is not truly longitudinal (i.e., we are comparing two separate samples rather than tracking the same individuals through time), the results show a clear pattern in that a higher propensity to plan corresponds to better wealth outcomes over time when the pandemic shock is considered. This result will be further explored in Section 6, along with the notion of economic resilience.

5.4 Financial literacy

The literature on the propensity to plan has long asserted the importance of financial literacy and how the two concepts are intertwined. Another advantage of the survey data is that there is information available on newspaper reading habits, and more specifically, readership of the Finance/Money section. We can test the impact of reading about finance on planning, using a distinct regression model. Given that there is almost certain simultaneity between literacy and planning, an IV model is specified using newspaper readership (i.e., reading any sections) as an instrument for financial Section readership. The second stage regression suggests a strong relationship between financial literacy and the propensity to plan: reading the Finance/Money Section of a newspaper would have a causal impact on propensity to plan of 13.6%. Of the other significant covariates, the Finance/Money dummy is by far the largest in terms of its impact on planning. The full results are in Appendix 5.

With the relationship between planning and literacy confirmed, a new model is specified using the initial IV model from Section 5 and adding the financial literacy variable (instrumented for newspaper reading) and with a modified planning dummy which takes the value 1 if the individual rates themselves as a 4 or a 5 (i.e., a 'higher planner) on the response scale for propensity to plan. This is, again, instrumented by the adherence to routine variable. The numerical similarity between the planning and financial literacy variables will facilitate interpretation. Table 19 shows the results of both regressions.

Table 19: Modified IV regression models with planning, financial literacy, and interaction terms

VARIABLES	(1) Basic IV model	(2) Interaction IV model
High Planner dummy	0.293***	
	(0.0621)	
High Literacy dummy	0.368***	
	(0.0354)	
High Planner × Low Literacy		0.348***
		(0.0853)
Low Planner × High Literacy		0.734***
		(0.214)
High Planner × High Literacy		0.546***
		(0.0673)
Single dummy	-0.226***	-0.225***
	(0.00892)	(0.00894)
Education	0.145***	0.146***
	(0.00252)	(0.00252)
Age	0.0581***	0.0584***
	(0.00161)	(0.00161)
Square of age ÷100	-0.0592***	-0.0595***
	(0.00183)	(0.00183)
Male dummy	0.377***	0.378***
	(0.00928)	(0.00941)
White dummy	0.125***	0.125***
	(0.0107)	(0.0107)
Rural	-0.146***	-0.146***
	(0.00876)	(0.00879)
Constant	-0.627***	-0.668***
	(0.0479)	(0.0622)
Observations	35,803	35,803
R-squared	0.279	0.278

Standard errors in parentheses. Significance of coefficients is shown as follows: *** p<0.01, ** p<0.05, * p<0.1. Results for the entire sample, less respondents who did not provide an answer. The planning variable is P1 (carefully considering options before a purchase decision). The dummies for are all binary variables that take 1 if the respondent is single, male, white, or rural, and 0 if not. Education is on a sliding scale of 1 (high school) to 7 (Master's or PhD completed), broadly corresponding to the notion of 'level' of education. Age is in years. See Appendix 7 for a full list of definitions. 'High' planners have self-reported planning at 4 or 5 out of 5, while 'high' literacy indicates that the respondent reads the Money/Finance Section of a newspaper. 'Low planners have self-reported planning at 1 or 2 out of 5, while 'high' literacy indicates that the respondent does not read the Money/Finance Section of a newspaper.

Much like the initial IV regression, model (1) shows a strong relationship between planning and wealth: being a high planner has a causal impact on the Wealth Index to the

tune of 18% compared to not being a high planner. As for financial literacy, the impact is even greater: individuals who read the Finance/Money Section of a newspaper have 23% higher wealth than those who do not. Model (2), which adds interaction terms to test the combined impact of planning and literacy on wealth, shows a much stronger impact of the dual skill of planning and literacy, compared with a baseline of low planning and low literacy or only having one without the other (while the Low Planner × High Literacy interaction variable has the largest coefficient, its standard error is significantly higher than the other two terms, and is therefore interpreted with caution). Overall, an individual who has both a high propensity to plan and a high level of literacy sees a wealth gain of 34%.

Discussion

The challenges of measuring wealth and identifying the causal impacts of planning are at the heart of this thesis. In this section, the analysis of both the Wealth Index and the regression results is provided, with respect to the literature and in terms of the implications of the findings. The limitations of each analysis will also be covered, along with avenues for further research.

Traditional or comprehensive? The future of measuring wealth

We constructed a novel indicator, the Wealth Index, that included income, homeownership, self-reported financial situation, and financial anxiety, which proved to be a successful measure of affluence in Canada. The results from age-Wealth Index profiles and the subsequent regressions models and robustness checks provided evidence that the index showed the expected relationships with appropriate predictor variables including age, education, gender, planning, ethnicity, living in a rural/urban area, risk, impatience, and ability. The empirical evidence for relationships between planning, financial literacy and wealth found by several studies was supported with a variety of results and several model specifications. The 'shadow analysis' Wealth Index also added confidence to the interpretations: even with a very different PCA approach, the results remained robust.

One key implication of this result is that novel measures of wealth are of analytical interest and can indeed shed new light on the way in which we view the economic well-being of the population as a whole. For example, contrary to the increase in financial net worth and GDP per capita in Canada observed between 2017 and 2021, the Wealth Index showed a modest overall increase of 2.7%, or an annualized increase of 0.67%. While this paints a wholly different portrait of wealth, it is easier to reconcile with certain reported realities of Canadians. Statistics Canada uses a 'Financial Resilience Index' to report on the level of financial stress across the country, which showed that Canadians' individual financial situations have deteriorated in many ways (Duncan & Koci, 2021). The proportion of Canadians who self-report their credit score as "good or better" declined from 91% in 2017 to 77% in 2021. In 2017, 64% of Canadians reported having a liquid savings buffer

of three months of more, which has trended downwards ever since: the current maximum estimate is 50%, with the lowest reaching 33%. These findings are not coherent with the idea that Canadians are becoming 'wealthier', and especially not to the tune of 16%, the reported increase in net worth between 2017 and 2021 (Statistics Canada, 2022). A more nuanced measurement and understanding of wealth would help to paint a more representative portrait of individuals' economic realities.

While the index performed reliably and consistently, it is challenging to vouch for its accuracy. In the absence of a 'true' value, there is no standard by which to compare the results of the index used in this thesis. Wealth indexes tend to be used in the developing world rather than in OECD countries (Filmer & Pritchett, 2001), so large-scale comparisons from reputable sources for Canadians are unavailable. The implication of this is that the Wealth Index should not be compared against existing measures of wealth, not least net worth, but rather viewed as a composite index whose objective is to join different facets of wealth and use them to test existing hypotheses in a novel way. The accuracy of the Wealth Index as a true measure is also a direct result of the variables used in the PCA: the richer and more varied the information, the more likely the index is to reflect objectively measurable levels of wealth. In this sense, a clear extension of this research is to run additional analyses with new or combined datasets.

The distribution of the Wealth Index does not match the distribution of gross assets across the Canadian population, whereby financial wealth is disproportionally concentrated in the richest few centiles. Given the variables included in the index, it might be more useful to think of the Wealth Index as the wealth potential index: individuals with higher values have greater capacity for accumulating more wealth, but that higher value does not imply that this potential has been realized. As an example, the Wealth Index includes a dummy variable for homeownership, but does not consider the type or the value of the property. In the index, the owner of a modestly valued property would be considered as 'equivalent' to an individual owning a property with a market value several times higher. The emphasis is on the potential of the homeowner to generate future wealth, rather than the amount of wealth. Contrary to the net worth approach, where the degree of variance at the lower end of the wealth distribution is relatively low, the use of the Wealth Index gives a more

granular level of detail to an 'average' individual, at the expense of that granularity at the higher end of the distribution. The result of this is that it can return insights that resonate more directly with 'on the ground' realities faced by populations.

Empirical support for the relationship between economic planning and wealth

The theoretical foundations and empirical evidence for the relationship between planning and wealth is recent: most of the literature on the propensity to plan has been written in the 21st century. The evidence from this literature is consistent with the results of this research insofar it shows a strong and demonstrable relationship, robust to different research techniques and measurements of planning and wealth. This research has confirmed Ameriks et al.'s (2003) United States-based findings in the Canadian context, with a larger sample size and a differently structured dataset. Regression results show that this relationship is highly relevant with respect to other determinants of wealth: a one-point increase in planning has an equivalent impact on wealth as one additional level of educational attainment (e.g., the difference between a bachelor's degree and a baster's degree) or three years of age (e.g., the additional wealth accumulation expected between one's 39th and 42nd birthdays). Planning appears to have a stronger effect on wealth than other cognitive and non-cognitive factors such as patience and risk-taking, which are well-documented determinants of affluence.

Several novel findings also emerge from the research undertaken for this thesis. Ameriks et al.'s (2003) results did not include covariates for race/ethnicity or for urban/rural settings. Both factors proved to be robust and consistent predictors of wealth: estimates suggested that being white increases wealth by 7.8%, compared with those reporting a different race/ethnicity, while living in a city does so by 9.8%, compared with residents of rural areas. Furthermore, more detailed regression results indicated that the planning-wealth relationship remained stable even when those covariates were added one by one, which implies that the effect of the propensity to plan is unaffected by ethnic background or living in an urban or rural area. Quantile regression results offered another valuable insight: that the impact of planning on wealth holds true across the wealth distribution (i.e., it is not a local or isolated relationship), but that is it much stronger in lower-wealth individuals. This result certainly merits more analysis in the future, given that relatively

little work has been done with respect to the effects of the propensity to plan at different levels of wealth. With respect to the Openness to experience, Conscientiousness, Extraversion, Agreeableness and Neuroticism (OCEAN) set of cognitive and non-cognitive skills, Mosca & McCrory (2016) reported findings that the importance of different traits varied across the wealth distribution, suggesting that there is more to be investigated on this front.

One key consideration in the interpretation of the results is related to the use of the instrumental variables. The assertion that adherence to a routine and planning activities for friends are suitable instruments for economic planning is fundamental to the conclusions that are presented. While we can reasonably have confidence in these results – given the robustness tests performed and the variety of specifications used – a future avenue for research is clearly to repeat the analysis with different instruments. In terms of relevance, the first-stage results showed that both instruments had a clear and significant impact on planning, providing reassurance that the first condition was met. In terms of the exogeneity condition, the assumption that routine adherence (z_1) and planning activities for friends (z_2) do not directly affect wealth cannot be tested as easily. It is conceivable that the second instrument does have a direct relationship with wealth accumulation because it implies extraversion, which has been identified as some cases as a determinant of affluence (Exley et al., 2022). The same could be true for routine, which might affect wealth through conscientiousness (i.e., discipline), another non-cognitive attribute. OLS models of the instruments on the wealth and planning variables, which showed extremely small coefficient estimates, gave some reassurance that the exclusion restriction criterion was met. That said, they are not guarantees and they are underpinned by the Gauss-Markov assumptions associated with OLS regression.

The implications of the findings are exciting for behavioural economics and personal finance, particularly because propensity to plan is a characteristic that is more obviously within individuals' scope of control. Unlike the other determinants of wealth considered – such as age, race/ethnicity, and gender – there are tangible steps that a motivated individual can take to increase both their understanding of what economic planning means (i.e., financial literacy, knowledge, and information about their own spending) and their

planning abilities (i.e., behaviours and habits that facilitate planning). The results suggest that the implications could be of particular relevance to young people and to lower-wealth individuals, in that they stand to improve their wealth outcomes by a greater degree. An individual wishing to acquire planning knowledge and skills can do so at a lower cost than attempting to change other attributes that determine of wealth. For example, acquiring additional education tends to be expensive and incurs an opportunity cost of wages. Moving from a rural area to an urban area in an attempt to increase wealth is a serious commitment, and could incur heavy social costs for, such as isolation or adaptation difficulties.

The policy implications of this relationship are wide-ranging. While there has been a great uptake of the idea of planning and financial literacy in the public sphere in recent years, there is space to better integrate the importance of planning into public education campaigns and policy decisions with respect to wealth redistribution. In many cases, these two concepts go conveniently hand in hand: for example, the Canadian government's creation and promotion of the tax-free first home savings account (FHSA) will likely improve both financial literacy outcomes (i.e., educating the public on the need to save in order to acquire a first property) and planning outcomes (i.e., incentivizing the action of planning regular deposits in a savings account towards the objective of homeownership). There is a clear link between the importance of the objective itself and the planning required to achieve it. However, in other cases, prior initiatives could be improved in order to better connect planning to literacy. In some cases, both planning and literacy initiatives have come under criticism, such as the introduction of the tax-free savings account (TFSA) in 2011. Few Canadians understood the modalities of this investment instrument, and the relationship between the TFSA and planning is not evident, in that the account simply reduces an individual's tax burden rather than contributing to a well-defined objective. Indeed, the TFSA has been criticized insofar as the name itself causes its misuse: rather than contribute gradually over time and let their savings grow, Canadians tend to use the accounts for short-term savings objectives, which is not the optimal use (CBC, 2016). In a behavioural experiment, Boyer et al. (2022) found that targeted interventions to improve individuals' understanding of TFSAs has a direct effect on the quality of contribution decisions. A more nuanced deployment of the TFSA would have

specifically encouraged medium- and long-term use in order for users to maximize their tax savings and wealth accumulation outcomes.

Beyond policy design, the growing attention to the propensity to plan also suggests that collecting data and deepening our collective understanding on the links between planning and wealth is a worthy social objective. The data used in this thesis is of a particularly high quality due to the large sample size and because it offers an opportunity to measure wealth and income outcomes and cognitive patterns within the same group of individuals. Other research could go even further with more standardized behavioural data, more precise questions about distinct types of planning abilities, and variables that do not rely on self-assessment. Outcomes of more granular and specific research would also give rise to more targeted policy direction and deepen our understanding of how diverse groups' relationships with wealth accumulation. In that context, the initial conclusions drawn from this study could be further tested and widened.

As shown by the results of Wealth Index, the impact of planning – a 'choice' variable – is strong and significant, but not drastic. A one-standard deviation (0.98 points) in planning increased wealth by 9.3% This contrasts with Ameriks et al.'s (2003) findings that implied up to a 40% impact on wealth (albeit with a different measure of wealth). Two implications for planning and literacy emerge from this fact. Firstly, the role of planning and literacy must not be overstated in determining individual outcomes: they are not in themselves sufficient to overcome other economic forces that affect wealth, and certainly not systemic ones. As argued by Pinto and Coulson (2011), governments and financial institutions can go further to recognize this reality. If this acknowledgment is absent, a tacit implication of promoting financial education can be that individuals have great or even full control over their level of economic well-being. Beyond its inaccuracy, this message is detrimental to our view of how wealth in unevenly distributed amongst the population. It is widely recognized as a positive development that banks and other financial institutions emphasize the importance of developing financial plans and following them: the evidence from this thesis suggests that this behaviour is beneficial both at the individual and at the societal level. Use of technology to this end is likely to encourage adoption of such behaviours, particularly among younger clients for whom the

relationship between propensity to plan and wealth accumulation is higher. Communities that are historically underserved by banks now have greater access to financial advice through the internet and smartphone applications.

Although banks are increasingly integrating financial literacy in their offerings, the Canadian financial services sector has seen cases in the past few years of institutions actively encouraging their clientele to use potentially damaging financial instruments. As is the case elsewhere in the world, banks have been identified as offering inappropriately high interest rates on loans, charging fees that are not suited to their clients' financial situations, and concealing key terms and conditions in contracts (CBC, 2021a; CBC, 2021b). The Government of Canada and provincial regulators have a role to play in holding institutions to account for such practices and fostering a financial services ecosystem in which a genuine 'financial literacy revolution' can take place.

Explaining the determinants of wealth

Age was identified early as a key determinant of wealth, and this research has provided evidence for a lifecycle hypothesis that deviates from neoclassical theory. While the original age-Wealth Index profiles show wealth accumulation increasing over time, the results of the Wealth Index suggest that an individual's earlier life is a more important moment for wealth accumulation. The implications of this result are important for how we view wealth and wealth inequality. It suggests that financial education could be more effective if their audience is younger, where, for the majority of the population, wealth-determining events occur. This includes decisions about education, and efforts to save in order to purchase wealth-generating assets. Additionally, the propensity to plan grows rapidly at a younger age, suggesting that this could be an ideal moment in an individual's lifecycle for targeted interventions to improve knowledge and skills around economic planning. Although the majority of tangible net worth generation takes place later in life (for the reasons described in the lifecycle hypothesis), this phenomenon only tends to be more relevant for high wealth individuals than lower wealth individuals.

Few surprises emerged from the other regression results. Two demographic factors associated with lower levels of wealth were being single (-14.8%) living in a rural area (-

9.6%); these relationships are well-known in the literature and specifically in Canada, where wealth has long been concentrated in highly concentrated urban areas. The impact of being single on wealth did increase after the onset of the pandemic: for the 2017 sample, the effect is -13.2%, while for the 2021 sample it is -16.6%. An intuitive, but unexplored, explanation for this could be that married individuals benefit from more financial security in the face of uncertainty because of having a dual income. The two demographic determinants with a positive impact on wealth were being a man and being white, two broadly recognized patterns that have been explored in other studies. These relationship between wealth and each of these two non-choice variables also became significantly stronger following the onset of the pandemic: for men, the impact went from +24.5% to +27.2%, while the impact of being white went from +6.5% to +9.3%. The results paint a pessimistic portrait on the effects of the shock on the gender and racial wealth gaps, which is coherent with early emerging (though also contested) literature about the disproportionate burden of the pandemic born by women and by people of colour (Power, 2020; Lim & Zabek, 2023). These impacts certainly constitute an avenue for further research.

As it relates to determinants of wealth, one key limitation of the regression results is that their explanatory and predictive power is only moderate: for the main IV models, the R-squared result was never higher than 0.3. While there is not an objectively 'good' R-squared, it is important to consider the conclusion that 70% of the variation in the Wealth Index is unexplained by the covariates used in the models. Several factors contribute to this. Firstly, certain well-known determinants of wealth are not available in the dataset, such as level of health or the socioeconomic status of an individual's parents. Secondly, some determinants of wealth are unobservable, for example, luck or talent. These results further support the argument in the previous sub-section that the role of planning and financial literacy should be contextualized in the idea that individuals' wealth outcomes are, for the most part, shaped by forces and factors outside of their scope of control. A further limitation is the arbitrary nature of the Wealth Index, which is expressed on a scale of 0 to 4.3, as determined by the input variables. Although percentage impacts can be deduced from regression, these cannot be translated into more concrete effects on monetary personal wealth with the data available.

Resilience and uncertainty

Resilience has long been connected to wealth in the macroeconomic sense. Governments across the developed world intervened drastically to support individuals during the COVID-19 pandemic, with measures ranging from direct stimulus payments (Quebec, the US), emergency unemployment benefit (Canada), targeted sectoral job support schemes (the UK), and several other fiscal initiatives. While clearly necessary at the onset of the pandemic and with near-total economic uncertainty regarding the months ahead, no economic consensus has been formed around the efficiency and effectiveness of these programs. The data analyzed in this thesis are not necessarily relevant to that debate. However, the results give support to the hypothesis that planning plays a strong role in wealth accumulation. From a policy perspective, individuals with higher economic resilience reduce the burden and pressure on governments to intervene following shocks and minimize the unintended effects of these interventions on the economy as a whole. From an individual's point of view, resilience contributes to greater autonomy and a better ability to weather financial shocks. Given the observation that individuals do not control the majority of their wealth outcomes, a policy challenge seems to emerge between the macroeconomic and microeconomic dimensions: how can governments create the conditions for resilience in order that individuals can better weather shocks when they occur? Initiatives such as the Government of Canada's Financial Literacy Strategy represent a step in the right direction. It will be important to take bold steps in implementing this vision and ensuring that it is done so in a way that is accessible to all of the population.

The results of the comparative analysis between the 2017 and 2021 samples showed that the impact of planning on wealth increased significantly after the onset of the pandemic. This result is corroborated by robustness checks, and by comparing the mean and median Wealth Index values for individuals of different levels of planning. In each case, the observation is clear: 'high planners' weathered the pandemic shock better than 'low planners' from a wealth perspective. While several reasons could underly this result, it seems reasonable to infer that planning contributes to the development of resilience. This is all the more likely given the observation that the distribution of Wealth Index determinants (household income, home ownership, financial situation, and financial

anxiety) does not change between 2017 and 2021, but rather the relative change of wealth level in similar groups. This insight would greatly benefit from a longitudinal study in which individuals were followed over several years. In particular, distinct groups could be established to determine the treatment effects of planning.

As it relates to resilience, a clear limitation of the model is that the Wealth Index is not sufficiently detailed to consider macroeconomic details, such as interest rate changes. While homeownership is taken as a contribution to wealth, the increases in the Bank of Canada's overnight rate since 2022 have led to increased financial pressure on homeowners, who are increasingly unable to cover mortgage payments comfortably, i.e., that all accommodation expenses constitute 30% of income (CMHC, 2022). This nuance is lost when homeownership is only included as a binary variable. With that being said, the inclusion of the financial situation and financial anxiety index variables does allow, to some extent, this sort of economic discomfort to be reflected in the Wealth Index. An interesting avenue for research in future datasets would be to compare the levels of self-reported financial well-being among homeowners before and after the consumption shock of rising interest rates. Due to lags in the transmission mechanism (e.g., fixed rate mortgages and capped variable rate mortgages), the full effect of this is likely to take several years before being felt by all property owners.

The mechanism linking economic planning and resilience feels intuitive: individuals who are more inclined to plan are probably more prepared for uncertainty. However, further empirical evidence would be required to support this statement more robustly. For this, more specific data would be required on the components of financial resilience, such as ability to manage debt and the amount of emergency savings an individual has available at a given moment. With these variables available, similar models could be specified to test whether engagement in planning has a tangible or financial impact on an individual's ability to cope with an economic shock. This is an example of where the methodology presented in this thesis could be joined to a more traditional quantitative approach in which 'hard' financial data is used (e.g., using 'balance of savings account' as a dependent variable). The collection of such information, while sensitive and expensive, could

provide valuable and nuanced insights as to our understanding of the causal mechanisms underlying planning and wealth.

6. Conclusion

The research carried out for this thesis has three key outcomes: a novel contribution to non-traditional measures of wealth; empirical support for the important role played by economic planning in determining individual wealth; and comparative results showing that individuals who plan fared significantly better in the context of the pandemic than those who do not. As a measure, the Wealth Index performed reliably and consistently when tested with long-accepted theories of wealth accumulation, notably the lifestyle hypothesis, the influence of non-choice variables such as gender and race, and the roles of the propensity to plan and financial literacy. The results that it generated were robust to different model specifications and when the index was generated in different ways.

Supporting results of previous empirical studies, this research demonstrated a significant and robust effect of planning on wealth, of comparable importance to other determinants of wealth. The analyses in this thesis also extend work already completed on the topic: using the concept of economic planning, adding new demographic controls such as geography and race/ethnicity, and showing that planning has an even higher relative impact for people with lower levels of wealth. Furthermore, the comparative analysis of the 2017 and 2021 samples in this research supports the idea that planning played a key role in individuals' levels of economic resilience, most likely by enabling them to better weather the economic uncertainty that resulted from pandemic.

Several avenues for future research emerge from the limits of this study. Given that only two reasonable candidates for instruments were available for the IV model, there is potential to further evaluate the robustness of the wealth-planning relationship with other instrumental variables. As it relates to planning and financial literacy, a dataset with more specific questions on economic behaviours (e.g., debt, credit card use) would allow for the conclusions drawn from this research to be more thoroughly tested. This is particularly true for understanding the role that resilience played in determining financial outcomes after the onset of the pandemic. Finally, the Wealth Index itself, and other similarly constructed indicators could be used to test (or retest) the numerous hypotheses from the literature regarding the determinants of individual wealth.

References

Alan, Sule (2006). "Precautionary wealth accumulation: Evidence from Canadian microdata." *Canadian Journal of Economics*, vol. 39, no. 4 p. 1105-1124.

Ameriks, John, Andrew Caplin, and John Leahy (2003). "Wealth Accumulation and the Propensity to Plan". *The Quarterly Journal of Economics*, vol. 118, no. 3, p. 1007-1048.

Ando, Albert, and Franco Modigliani (1963). "The "life cycle" hypothesis of saving: Aggregate implications and tests." The American Economic Review vol. 53 no.1, p. 55-84.

Arrow, Kenneth J., Partha Dasgupta, Lawrence H. Goulder, Kevin J. Mumford, and Kirsten Oleson. (2012). "Sustainability and the measurement of wealth." *Environment and development economics*, vol. 17, no. 3, p. 317-353.

Blanchard, Robert D., James B. Bunker, and Martin Wachs (1977). "Distinguishing aging, period and cohort effects in longitudinal studies of elderly populations." Socio-Economic Planning Sciences 11, vol. 3, p. 137-146.

Block, Sheila, Grace-Edward Galabuzi, Ricardo Tranjan (2019). "Canada's Colour Coded Income Inequality". Canadian Centre for Policy Alternatives.

Boyer, M. M., d'Astous, P., & Michaud, P. C. (2022). Tax-Preferred Savings Vehicles: Can Financial Education Improve Asset Location Decisions?. *Review of Economics and Statistics*, 104(3), 541-556.

Bucci, Alberto, Riccardo Calcagno, Simone Marsiglio, and Tiago Neves Sequeira. "Financial literacy, human capital and long-run economic growth." CeRP, Center for Research on Pensions and Welfare Policies, 2023.

Bütler, Monika (2001). "Neoclassical life-cycle consumption: A textbook example." Economic Theory vol. 17, p. 209-221.

CBC (2016, January 2). “Tax-free savings accounts a good product with a bad name, experts say”. Retrieved from https://www.cbc.ca/news/business/rrsp/tfsa-confusion-rrsp-season-1.3369742

CBC (2021, January 30). “'Canadians deserve better': Experts decry 'outrageous' interest rates by alternative lenders”. Retrieved from: https://www.cbc.ca/news/canada/alternative-lenders-marketplace-1.5891676

CBC (2021, February 11). “Employees' claims of sales pressure spark shareholder lawsuit against TD Bank”. Retrieved from: https://www.cbc.ca/news/canada/td-class-action-upselling-1.5909176

CPA Canada (2022). “Surviving or Thriving Study 2022”. Retrieved from https://www.cpacanada.ca/thrive

Dawson, Chris (2023). Looking on the (B)right Side of Life: Cognitive Ability and Miscalibrated Financial Expectations. Personality and Social Psychology Bulletin, 01461672231209400.

Di Matteo, Livio (1997). "The determinants of the wealth and asset holding in nineteenth-century Canada: Evidence from microdata." *The Journal of Economic History* vol. 57, no. 4 p. 907-934.

Di Matteo, Livio (2016). "Wealth distribution and the Canadian middle class: Historical evidence and policy implications." *Canadian Public Policy*, vol. 42, no. 2 p. 132-151.

Duncan, Eloise and Koci, Kujtim (2021, November). “The financial resilience and financial well-being of Canadians with low incomes: insight and analysis to support the financial empowerment sector”. Seymour Management Consulting Inc.

Exley, Jim., Doyle, Patrick. C., Grable, John., & Campbell, W. Keith (2022). “OCEAN wealth profiles: A latent profile analysis of personality traits and financial outcomes.” *Personality and Individual Differences*, vol. 185, 111300.

Filmer, Deon, and Lant H. Pritchett (2001). "Estimating wealth effects without expenditure data—or tears: an application to educational enrollments in states of India." Demography 38, vol. 1, p. 115-132.

Financial Consumer Agency of Canada (2021). “Make Change that Counts: National Financial Literacy Strategy 2021-2026”. Retrieved from https://www.canada.ca/content/dam/fcac-acfc/documents/programs/financial-literacy/financial-literacy-strategy-2021-2026.pdf

Financial Consumer Agency of Canada (2022). Financial Literacy Month 2022 Kick-Off. Retrieved from https://www.canada.ca/en/financial-consumer-agency/news/2022/10/financial-literacy-month-2022-kick-off0.html

Gathergood, J., & Disney, R. F. (2011). “Financial literacy and indebtedness: New evidence for UK consumers”. Available at SSRN 1851343.

Gruijters, Rob J., Zachary Van Winkle, and Anette E. Fasang (2023). "Life Course Trajectories and Wealth Accumulation in the United States: Comparing Late Baby Boomers and Early Millennials." *American Journal of Sociology*, vol. 129 no. 2, p. 530-569.

Harding, John P., and Stuart S. Rosenthal (2017). "Homeownership, housing capital gains and self-employment." *Journal of Urban Economics*, vol. 99, p. 120-135.

Huston, S. J. (2010). Measuring financial literacy. *Journal of Consumer Affairs*, vol. 44 no. 2, p. 296-316.

Juster, F. Thomas, Smith, James, & Stafford, Frank (1999). The measurement and structure of household wealth. *Labour Economics*, vol. 6 no. 2, p. 253-275.

Karlan, D., & Morduch, J. (2010). “Access to finance”. In *Handbook of Development Economics*, vol. 5, p. 4703-4784. Elsevier.

Kennickell, Arthur B., and R. Louise Woodburn (1999). "Consistent Weight Design for the 1989, 1992 and 1995 SCFs, and the Distribution of Wealth." *Review of Income and Wealth* vol. 45, no. 2 p. 193-215.

Letkiewicz, Jodi, Chris Robinson, Dale A. Domian, Natallia Uborceva (2015). “Behavioral and Wealth Considerations for Seeking Professional Financial Planning Help.

Lim, Katherine, & Zabek, Mike (2023). “Women’s labor force exits during covid-19: Differences by motherhood, race, and ethnicity.” *Journal of Family and Economic Issues*, p. 1-24.

Lusardi, Annamaria, and Olivia S. Mitchell (2008). "Planning and financial literacy: How do women fare?." *American Economic Review*, vol. 98, no. 2, p. 413-417.

Lusardi, Annamaria and Olivia S. Mitchell (2014). “The Economic Importance of Financial Literacy: Theory and Evidence”. *Journal of Economic Literature*, vol. 52, no. 1, p. 5-44.

Lusardi, Annamaria, and Peter Tufano (2015). "Debt literacy, financial experiences, and overindebtedness." *Journal of Pension Economics & Finance*, vol. 14, no. 4, p. 332-368.

Lusardi, Annamaria, Pierre-Carl Michaud, and Olivia S. Mitchell (2017). "Optimal financial knowledge and wealth inequality." *Journal of Political Economy*, vol. 125, no. 2, p. 431-477.

Lusardi, Annamaria, and Olivia S. Mitchell (2017). "How ordinary consumers make complex economic decisions: Financial literacy and retirement readiness." *Quarterly Journal of Finance* vol. 7, no. 3.

Micklewright, John et Schnepf, Sylke V. “How reliable are income data collected with a single question?.” *Journal of the Royal Statistical Society*, Series A: Statistics in Society, 2010, vol. 173, no 2, p. 409-429.

Mosca, Irene, & Cathal McCrory (2016). “Personality and wealth accumulation among older couples: Do dispositional characteristics pay dividends?” *Journal of Economic Psychology*, vol. 56, p. 1-19.

Nienhuis, C. P., & Lesser, I. A. (2020). The Impact of COVID-19 on Women's Physical Activity Behavior and Mental Well-Being. International journal of environmental research and public health, 17(23), 9036.

OECD (2022). “Gross domestic product (GDP): GDP per capita, USD, current prices and PPPS”. Retrieved from https://stats.oecd.org/index.aspx?queryid=61433

OECD (2013). “Measuring household wealth through surveys”. OECD Guidelines for Micro Statistics on Household Wealth. Retrieved from https://www.oecd-ilibrary.org/measuring-household-wealth-through-surveys_5k46dbqln17d.pdf?itemId=%2Fcontent%2Fcomponent%2F9789264194878-9-en&mimeType=pdf

O’Neill, Barbara, Jing J. Xiao, and K. Ensle. "Propensity to plan: A key to health and wealth." *Journal of Financial Planning*, vol. 29, no. 3 (2016): 42-50.

Pew Research Center (2019, February 14). “Millennial life: How young adulthood today compares with prior generations”. Retrieved from https://www.pewresearch.org/social-trends/2019/02/14/millennial-life-how-young-adulthood-today-compares-with-prior-generations-2/

Pinto, Laura Elizabeth, and Elizabeth Coulson (2011). "Social justice and the gender politics of financial literacy education." *Journal of the Canadian Association for Curriculum Studies*, vol. 9, no. 2 p. 54-85.

Power, Kate (2020). The COVID-19 pandemic has increased the care burden of women and families. *Sustainability: Science, Practice and Policy*, vol. 16, no. 1, p. 67-73.

Rutstein, S. O. (2015). “Steps to constructing the new DHS Wealth Index.” *Rockville, MD: ICF International*, vol. 542.

Santini, Fernando De Oliveira, Wagner Junior Ladeira, Frederike Monika Budiner Mette, Mateus Canniatti Ponchio (2019). “The antecedents and consequences of financial literacy: a meta-analysis.” *International Journal of Bank Marketing*, vol. 37, no. 6, p. 1462-1479.

Scholz, John Karl, Ananth Seshadri, and Surachai Khitatrakun (2006). "Are Americans saving “optimally” for retirement?." *Journal of Political Economy*, vol. 114, no.4, p. 607-643.

Statistics Canada (2020). Survey of Financial Security (SFS). Retrieved from https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getInstanceList&Id=1252634

Statistics Canada (2023). Canadian Income Survey, 2021. Retrieved from https://www150.statcan.gc.ca/n1/daily-quotidien/230502/dq230502a-eng.htm

Statistics Canada (2023). “Racialized Canadians are less likely to find as good jobs as their non-racialized and non-Indigenous counterparts early in their careers”. Retrieved from https://www150.statcan.gc.ca/n1/daily-quotidien/230118/dq230118b-eng.htm

Statistics Canada (2022). Distributions of household economic accounts for wealth of Canadian households, second quarter 2022. Retrieved from The Daily — Distributions of household economic accounts for wealth of Canadian households, second quarter 2022 (statcan.gc.ca)

Statistics Canada (2022). Employment growth in Canada and the United States during the recovery from COVID-19. Retrieved from https://www150.statcan.gc.ca/n1/pub/36-28-0001/2022012/article/00001-eng.htm

Statistics Canada (2022). To buy or to rent: The housing market continues to be reshaped by several factors as Canadians search for an affordable place to call home. Retrieved from https://www150.statcan.gc.ca/n1/daily-quotidien/220921/dq220921b-eng.htm

Statistics Canada (2020). Survey of Financial Security (SFS). Retrieved from https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=1252634

Stock, James H., Jonathan H. Wright and Motohiro Yogo (2002). A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments. *Journal of Business & Economic Statistics* vol. 20, no.4, p.518-539.

Walker, Brian, Leonie Pearson, Michael Harris, Karl-Göran Maler, Chuan-Zhong Li, Reinette Biggs, and Tim Baynes (2010). "Incorporating resilience in the assessment of inclusive wealth: an example from South East Australia." *Environmental and Resource Economics*, vol. 45, p. 183-202.

Wolla, Scott A., and Jessica Sullivan (2017). "Education, income, and wealth." Page One Economics®. Retrieved from https://research.stlouisfed.org/publications/page1-econ/2017/01/03/education-income-and-wealth

Appendices

Appendix 1: Lifestyle Survey Privacy Notice

[The company] and its affiliates ("our", "us" and "we") is a global a digital and media advertising company. Our core business is to help our clients improve how they advertise and market, whether by print, post, email or on websites. Information about people like you is therefore important to our business.

This privacy notice ("Notice") tells you more about the information – specifically, the personal data - that we gather through the Lifestyle Survey and any follow-up surveys (together, the "Survey").

The Survey is our consumer survey, which provides insight into the types of people that might buy our clients' products or services. Thanks to the information that you share when completing the Survey, you have the power to influence some of the world's largest brands.

In brief, this Notice explains:

why we process your personal data

the legal grounds that allow us to process your personal data;

where the personal data comes from;

the types of personal data we process;

who gets to see the data;

how long we keep the data;

how we protect your personal data;

your rights in relation to your personal data; and

how to contact us.

In the Supplementary Information Section of this notice, we explain what is meant by "personal data", "processing" and other terms used in the Notice.

1. Why we process your personal data

We never use personal data collected through the Survey to target you with our clients' marketing activities.

Personal data provided through the Survey is used by us for research purposes – to allow us to better understand how consumers respond to marketing and advertising. Your personal data can help us decide the right types of people to target when carrying out marketing and advertising activities on behalf of our clients. To do this we use the personal data described in Section 4 to gain audience insight for multiple clients.

We may also use the personal data you provide through the Survey to find individuals that behave in a similar manner to you and to then send those similar individuals targeted marketing and advertising on behalf of our clients. This type of activity may include what's known as "profiling" (see Profiling and Your Right to Object to Profiling). As previously mentioned, we do not target people who have taken part in the Survey.

2. What are the legal grounds that allow us to process your data?

Data protection law allows us to process your personal data if we have a legitimate interest in doing so. As described in Section 1, we use your personal data for research purposes: to derive insight into the audiences that are relevant to our clients. We will not process your personal data if your interests, rights and freedoms override our own interests.

3. Where the personal data comes from

We employ market research companies to run the Survey. You will have voluntarily registered with one or more of these market research companies.

4. Types of personal data we process

Each time you complete the Survey through a market research company, we receive a copy of your Survey responses. To help with our analysis of your Survey responses, your responses are linked to a unique reference code (a random arrangement of numbers and letters). This reference code is given to you by the market research company.

We do not receive your contact details (such as your name or email address) from the market research companies. This type of directly identifiable data is removed from the data set before being shared with us.

The Survey asks for your opinions, interests, hobbies; views on products and services items you own and buy; your age; occupation; income; lifestyle; media usage; views on marketing and advertising; and other information that is useful for describing different types of people.
Where legally permitted, we may ask about your sexual orientation or your religious beliefs, your race or ethnicity. It is always up to you whether you provide this information when completing the Survey. We never use this information to target you with marketing and advertising.
COOKIES.
When you complete the Survey online, a cookie may be set and stored in your browser. The market research company will ask your permission to set the cookie. The cookie records your unique reference code and can be synced with other cookies on your device – for instance cookies that record the websites you have visited. This allows us to find other people that behave in a similar way to you. As previously stated, we only use this information for research purposes; never to target you with marketing and advertising.
PROFILING.
Some of what we do with your personal data involves what's known as "profiling" – using automated means to process your personal data to analyse or predict your personal preferences, interests, or behaviour. This helps with our research – it enables us to work out how best to target a person with the same preferences, interests or behaviours as you.
Please note that we do not carry out profiling activities in order to target you with our clients' marketing and advertising activities. You will not receive targeted marketing as a consequence of the profiling activity described in this Notice. Instead, profiling that involves your personal data enables us to derive insight into our clients' target audiences so that we can determine when, where and how to market to those audiences. You can object to our profiling (see Section 5 of this Notice, below).

5. Who gets to see your personal data?

Most of the analysis of the personal data described at Section 4 of this Notice uses data from aggregated groups of Survey respondents ("Survey Results"), not individual responses. We share Survey Results, not individual data, with [company] staff responsible for creating consumer insight reports and media plan recommendations for our clients. We also share Survey Results with our clients.
We share individual responses to the Survey with: (i) a small team within [the company]; (ii) the company that provides us with cross tabulation software – that's software which enables us to carry out faster analysis; and (iii) a client of [the company] where the client has commissioned a specific Survey.

6. How long do we keep your personal data?

We keep personal data for the length of time it takes us to pursue our legitimate interest, which is explained in Section 2. When we no longer need your personal data, the data is securely deleted.

7. How We Protect Your Personal data

Our safeguards include robust systems and processes designed to ensure that we collect only the minimum personal data necessary to pursue our legitimate interest, and that only those who need to view your personal data can see it. We carry out checks to make sure we adhere to restrictions on our use of the information (such as making sure that no marketing or advertising is directed at individuals who have provided data through the Survey).
We have implemented policies, processes and systems to help keep your personal data secure. Our Global Information Security Program is based on the ISO27001/2 standards, and industry and internal best practices. We continuously monitor and improve standards and regularly test our security measures. We also maintain an incident response plan for dealing with any incident or breach where your information may be put at risk or compromised, including measures for logging and audit trails, incident detection and security incident information gathering and reporting.

8. Your rights

If you are from the European Economic Area, you have rights (with some exceptions and restrictions) to:
object to our processing of your personal data, including profiling. You can object, on grounds relating to your particular situation, at any time. In which case, we shall stop processing the data that your objection relates to, unless we can show compelling legitimate grounds to continue that processing;
access your personal data. If you make this kind of request and we hold personal data about you, we are required to provide you with information on it, including a description and copy of the personal data and why we are processing it
request erasure of your personal data in certain circumstances;

request correction or updating of the personal data that we hold about you and that is inaccurate;
request the restriction of our processing of your personal data in some situations. If you request this, we can continue to store your personal data but are restricted from processing it while the restriction is in place;
complain to your local data protection authority about our collection or use of your personal data. For example, in the UK, the local data protection authority is the UK Information Commissioner's Office.
Please note the limited data that we receive about you, as described in Section 4 of this Notice. Therefore, if you choose to exercise the rights described above, we may ask you to provide additional information so that we can satisfy ourselves as to your identity before we take further action.
If you are from the European Economic Area and would like to exercise any of these rights in relation to any information that we hold about you, please contact us. Our contact details can be found in Section 10 of this Notice. We will consider and respond to your request in accordance with the relevant law.

9. Transferring personal data from and to Europe

Your personal data is kept in on servers in multiple countries and may be transferred outside of the European Economic Area for analysis. We may transfer personal data to countries where our overseas group companies or clients are located to allow them to review the research and insight.
When we make any of these transfers, we take appropriate steps to ensure EU data protection law is complied with. These steps might include, for example, transferring the information to someone in a country which the European Commission has decided provides adequate protection for personal data, or to someone who has signed standard contractual clauses approved by the European Commission, our intra-group agreement includes these standard contractual clauses, to cover transfers to our non-European affiliates.

10. How to contact us

When processing your personal data in the ways described in this Notice, we act as a data controller. If you have any questions about this Notice or would like to exercise any of the rights mentioned in Section 10 this Notice, you can contact our Data Protection Officer in any of the following ways:

11. Changes to the Notice

We may make changes to this Notice on occasion. We will post any revised versions of this Notice on our website. Please review this Notice periodically to see if any changes have been made.

SUPPLEMENTARY INFORMATION

In this Supplementary Information section, we explain some of terminology used in the Notice.
"controller" – the person or company that controls the purposes and means of processing personal data.
"European Economic Area" – the 28 countries in the European Union plus Iceland, Liechtenstein and Norway.
"personal data" – any information that relates to you (or from which you can be identified).
"processing" – doing anything with personal data. For example, collecting it, storing it, disclosing it and deleting it.
"profiling" – using automated means to process personal data in order to work out certain things about people, like analysing or predicting their performance at work, reliability, economic situation, personal preferences, interests, behaviour, location or movements.
"transfer" – sending personal data outside the European Economic Area (e.g. by storing it on equipment located outside the European Economic Area), or allowing someone from outside the European Economic Area to access the personal data.

Appendix 2: Full PCA results

2.1 PCA results, sorted by factor loading score on Comp3

VARIABLES	Comp1	Comp2	Comp3	Comp4	Comp5
Household income level	**0,0065**	**-0,0111**	**0,2032**	**-0,0386**	**0,1375**
Household income level (broad)	0,0059	-0,0123	0,1832	-0,0351	0,1253
Age	-0,0523	-0,0001	0,1812	-0,0486	-0,0218
Personal income level	0,0001	0,0306	0,1668	-0,0464	0,1815
Personal income level (broad)	-0,0008	0,03	0,1556	-0,0477	0,1753
Own house/apartment	**-0,0092**	**-0,007**	**0,1483**	**-0,0593**	**0,1186**
Relationship status	-0,0164	0,0073	0,1166	-0,0401	0,0437
Travel by air for holiday	0,0793	-0,036	0,1155	-0,0027	0,0286
Travel by air for business	0,0785	-0,0657	0,113	0,022	0,0559
Financial situation	**0,0386**	**-0,0129**	**0,112**	**-0,0466**	**0,0552**
Air travel for holiday	0,0574	-0,0184	0,1032	-0,017	0,041
Air travel for business	0,0747	-0,0673	0,1019	0,0176	0,0528
Educational level reached	0,0233	0,0005	0,1	-0,0223	0,013
National daily - read in last year	0,0356	0,0277	0,099	0,0426	-0,0164
Live with partner/spouse	-0,0112	0,0282	0,0966	-0,0353	0,1096
Magazines - monthly	-0,0016	0,0838	0,0941	0,017	-0,0244
Work full time	0,0061	-0,0092	0,0923	-0,0269	0,0994
Time per week - read newspapers	0,0705	-0,0321	0,0905	0,0794	-0,0315
Staycation - Do regularly	0,021	0,0644	0,0863	0,0241	-0,014
I am happy with my standard of living	0,045	0,0277	0,0846	-0,0794	0,0266
DIY projects - Do regularly	0,0175	0,0432	0,0826	0,0137	0,0104
Use tablet	-0,0051	0,0596	0,0825	0,0165	0,087
Gardening - Do regularly	0,0107	0,0437	0,0824	0,006	-0,0397
Eat well-balanced meals	0,0504	0,0587	0,0821	-0,0988	-0,0927
Gender	0,0222	-0,0686	0,0783	0,0318	0,0693
Internet activities	0,0538	-0,0102	0,0778	-0,0011	0,1018
Time per week - newspaper websites/apps	0,0557	-0,011	0,0762	0,0607	-0,0109
I regularly exercise	0,0629	0,0338	0,0734	-0,0653	-0,0546
Time per week - printed newspapers	0,0609	-0,0435	0,0731	0,0709	-0,0426
Daily newspaper - read in last year	0,0107	0,0454	0,0724	0,0326	-0,0205
Soccer - Regularly do	0,0144	0,0115	0,0722	0,0234	0,0412
Motivations behind newspaper reading - to be informed	-0,0347	0,1114	0,0719	0,0234	-0,0273
Travel by bicycle	0,0785	-0,0654	0,0718	0,0374	-0,0606
Time spent on bike last time used	0,0621	-0,0404	0,0713	0,0556	-0,0566
Motivations behind magazine reading - to be informed	-0,0002	0,0768	0,0695	0,0372	-0,0227
International daily - read in last year	0,0669	-0,0111	0,0677	0,0402	-0,022

Have tablet	-0,0156	0,0793	0,0655	0,0221	0,1112
Motivations behind social media use - do not do this activity	-0,0352	-0,0315	0,0645	-0,0174	-0,0215
Internet activities	0,0437	-0,0031	0,0635	0,0059	0,106
Cycling - Regularly do	0,0213	0,0165	0,0631	0,0239	-0,0491
Exhibitions/galleries/museums - Do regularly	0,0389	0,0293	0,0621	0,0622	-0,0644
Chief income earner	0,0177	-0,0597	0,0596	0,0053	-0,0214
Motivations behind radio listening - to be informed	-0,0061	0,0634	0,0588	0,0262	-0,0058
Sport very important	0,0746	-0,0106	0,0584	-0,0265	0,0385
Recycle	0,0124	0,0744	0,0581	-0,0605	-0,0935
Time per week - read magazines	0,0865	-0,0554	0,0576	0,0882	-0,0307
Charity/Volunteer work - Do regularly	0,0292	0,0241	0,057	0,0257	-0,0576
Buy local	0,0543	0,0466	0,0567	-0,0843	-0,0952
Have Blu-ray DVD player	-0,0194	0,0695	0,0566	0,0437	0,0487
Have smart home utility manager	0,0336	0,0066	0,0566	0,0385	0,0543
Time per week - printed magazines	0,071	-0,055	0,0564	0,0757	-0,0352
Use digital radio	0,0208	0,0265	0,0564	0,0522	0,0608
Go to classical music concerts/opera - Do regularly	0,051	-0,0025	0,056	0,0485	-0,0488
Use smart home utility manager	0,0288	0,0154	0,0544	0,0458	0,0514
Travel by car	-0,0181	0,0617	0,0541	-0,0411	0,1251
Have digital radio	0,0223	0,0316	0,0533	0,0374	0,0658
Have eReader	0,0022	0,05	0,0528	0,0403	0,0409
Use Blu-ray DVD player	-0,0099	0,0485	0,0519	0,0533	0,024
Use Blackberry	0,0078	-0,0104	0,0506	0,009	0,0039
Motivations behind TV watching - to be informed	0,0008	0,0513	0,0505	0,0366	-0,02
Films - History	0,0159	0,0317	0,0493	0,0616	-0,0468
Time per week - travel by bicycle	0,0601	-0,0571	0,049	0,0717	-0,059
Films - Foreign (dubbed)	0,031	0,0101	0,0481	0,0571	-0,0844
Use eReader	0,0089	0,0422	0,0481	0,0458	0,0227
Travel by walking	0,0235	0,0458	0,0474	0,0239	-0,0757
Sailing - Regularly do	0,0389	-0,0023	0,0453	0,0329	-0,0018
Working out at gym - Regularly do	0,0217	-0,0131	0,0435	0,0475	0,0182
Actively involved in local community issues	0,1012	-0,0109	0,0434	-0,0564	-0,0498
Motivations behind newspaper reading - to relax/unwind	0,0442	-0,0062	0,043	0,0372	-0,0114
Social life satisfaction	0,0588	0,0415	0,0428	-0,0748	-0,0168
Motivations behind magazine reading - to relax/unwind	0,016	0,0579	0,0415	0,0312	-0,0293
Will buy smart home utility manager	0,0362	0,0142	0,0414	0,0441	0,0342
Time per week - magazine websites/apps	0,0779	-0,0402	0,0413	0,0731	-0,018
Restaurants non-fast food - Do regularly	0,0064	0,0975	0,0412	0,0377	0,0031
Surfing/Windsurfing - Regularly do	-0,0092	0,0896	0,0412	0,0139	-0,0645
Films - Documentary	0,0223	0,0306	0,0411	0,0532	-0,0679

Theatre - Do regularly	0,0383	0,0381	0,0401	0,0614	-0,0305
Motorbike racing - Regularly do	0,0311	0,0433	0,04	0,0296	-0,0223
Motor racing - Regularly do	0,0284	-0,0101	0,0388	0,0452	-0,0168
Try to go to different places	0,0637	0,0409	0,0378	-0,0665	-0,033
Look out for healthy products	0,0755	0,0576	0,0378	-0,0958	-0,0848
Public Sector/Government	-0,02	0,0106	0,0368	-0,0076	0,0118
Am seen as a positive person	0,0459	0,0779	0,0368	-0,0782	-0,0398
Dancing - Regularly do	0,0179	0,0041	0,0366	0,0299	0,0094
Hiking - Regularly do	0,0228	0,008	0,036	0,0362	-0,0154
Gym/Health club/Leisure centre - visits per month	0,0252	0,051	0,0352	-0,0031	-0,0011
Use iPod or other music player	0,0093	0,0571	0,0344	0,0634	-0,0017
Important to look after environment and take care of nature	0,0529	0,0605	0,0344	-0,0623	-0,1334
Magazines - fortnightly	0,0648	-0,055	0,0343	0,0533	-0,018
Use smart home appliances	0,0478	-0,0065	0,0341	0,0559	0,02
Spend time with family - Do regularly	-0,0243	0,1131	0,0341	0,0119	0,0131
Time spent on train last time used	0,0734	-0,0667	0,0337	0,072	-0,0765
Time spent on metro last time used	0,0803	-0,0869	0,0337	0,0619	-0,0762
Magazines - weekly	0,0613	-0,0049	0,0337	0,0274	-0,0129
Use streaming device	0,0103	0,0542	0,0335	0,0511	0,0838
Use intelligent home assistant	0,0436	-0,0165	0,0325	0,0633	0,0265
Arts/Cultural venues - visits per month	0,0617	-0,0232	0,0324	0,0466	-0,0475
Will buy smart home appliances	0,0443	0,0007	0,0324	0,044	0,0243
Will buy intelligent home assistant	0,0396	0,0012	0,0321	0,0538	0,0404
I often put my phone away so that I can really enjoy what I am doing	0,0583	-0,0313	0,0321	-0,0328	-0,0666
Like to learn about art and culture	0,0627	0,0635	0,032	-0,0172	-0,1068
Use laptop/desktop	-0,041	0,0932	0,0317	0,0367	-0,0084
Visit place of worship - Do regularly	0,0312	0,0051	0,031	0,0231	-0,0303
Have streaming device	0,0127	0,0592	0,0309	0,0525	0,1009
Motivations behind radio listening - to relax/unwind	0,0006	0,0589	0,0309	0,0284	-0,019
On the look-out for healthier alternatives	0,0689	0,0685	0,0306	-0,0817	-0,0745
Have iPod or other music player	-0,0053	0,0829	0,0304	0,0618	0,0391
Golf - Regularly do	0,0267	0,0471	0,0304	0,018	-0,0071
Prefer family over friends	0,0343	0,023	0,0304	-0,0621	0,0401
Martial arts - Regularly do	0,0234	-0,0185	0,0302	0,044	0,0112
Extreme sports - Regularly do	0,029	-0,011	0,0302	0,0403	0,0044
Time spent on motorbike last time used	0,0768	-0,1032	0,03	0,0538	-0,022
Motivations behind newspaper reading - to be entertained	0,0296	0,0242	0,0299	0,0461	-0,0282
Films - Arthouse	0,0375	-0,0125	0,0294	0,0564	-0,0749
Jogging/Running - Regularly do	0,025	-0,0143	0,0291	0,0505	-0,0048
I have more ability than most people	0,0679	0,0277	0,0291	-0,05	0,0122

Travel by motorbike	0,0912	-0,1199	0,0287	0,0438	-0,0141
Time per week - travel by car	0,0084	0,0322	0,0285	0,0094	0,0635
Supermarket - visits per month	-0,0167	0,0723	0,0285	0,0041	0,0006
Use wearable fitness device	0,014	0,0588	0,0284	0,0458	0,0772
I like to lead others	0,0786	0,0414	0,0284	-0,049	0,01
Have intelligent home assistant	0,0484	-0,0242	0,0282	0,0561	0,0276
Used branded mobile app to pay recently	-0,0053	0,0665	0,0281	0,0365	0,0218
Time per week - travel by road	0,0557	-0,0079	0,0279	0,0815	-0,0368
Time spent in car last time used	0,03	-0,0013	0,0277	0,0344	0,0241
Motivations behind newspaper reading - to be inspired	0,0447	0,0142	0,0276	0,0468	-0,0304
Prefer Canadian brands	0,0321	0,0555	0,0276	-0,0708	-0,0579
Football grounds/Stadiums - visits per month	0,0556	-0,0397	0,0274	0,0302	0,01
Fun at parties	0,0687	0,0424	0,0274	-0,0745	-0,0119
Have smart home appliances	0,0511	0,0003	0,0271	0,0454	0,0275
Have wearable fitness device	0,008	0,0702	0,027	0,042	0,0913
Days out - Do regularly	0,0254	0,0762	0,0264	0,0458	-0,0099
Motivations behind magazine reading - to be entertained	0,0142	0,0689	0,026	0,0435	-0,02
Live with children or siblings under 18	0,0062	0,0175	0,0259	-0,0107	0,0918
Important to do something for the good of society	0,0734	0,0636	0,0258	-0,0671	-0,0997
Time spent walking last time done	0,0338	-0,0063	0,0256	0,0533	-0,07
Good food and drinks are important	0,0509	0,0966	0,0255	-0,0747	-0,063
Prefer off-the-beaten-track holidays	0,0657	0,0156	0,0254	-0,0501	-0,0886
Squash - Regularly do	0,0317	-0,0253	0,0252	0,0529	-0,0059
Have laptop/desktop	-0,0416	0,1007	0,0252	0,0343	0,0155
Science/Engineering/IT	0,0163	-0,0247	0,0248	0,0168	0,0464
Skiing/Snowboarding - Regularly do	0,0134	0,056	0,0248	0,0045	-0,0041
I pay more for good quality	0,0675	0,0462	0,0248	-0,0486	0,033
Separated/Divorced/Widowed	-0,0185	-0,0014	0,0242	-0,0069	-0,0578
Challenge myself and be the best I can	0,0727	0,08	0,024	-0,0841	-0,0551
Athletics (track and field) - Regularly do	0,0385	-0,0166	0,0237	0,0484	-0,0055
Will buy tablet	0,0252	0,0067	0,0237	0,0329	0,0127
Manufacturing/Construction	-0,0005	-0,007	0,0237	-0,0135	0,0236
Will buy digital radio	0,0394	-0,0219	0,0235	0,0585	0,0023
Cricket - Regularly do	0,021	-0,0087	0,0235	0,0442	-0,0184
Will buy latest gaming handheld	0,0468	-0,0113	0,0234	0,0688	0,0054
Cooking/Baking - Do regularly	-0,0062	0,0977	0,0228	0,0148	-0,058
Motivations behind radio listening - to be entertained	-0,0025	0,0723	0,0222	0,0245	0,0034
Time per week - travel by motorbike	0,0547	-0,0715	0,0221	0,0524	-0,0146
Pop/rock concerts or festivals - Do regularly	0,0349	0,0409	0,0218	0,0585	-0,0246
Basketball - Regularly do	0,0373	0,0026	0,0217	0,0514	0,0207

Motivations behind magazine reading - to be inspired	0,023	0,0561	0,0217	0,0259	-0,054
Cinema - Do regularly	0,021	0,0702	0,0213	0,0673	-0,0233
Important to learn from other cultures	0,0511	0,0786	0,0212	-0,041	-0,1048
Films - Drama	-0,0046	0,0707	0,0209	0,0376	-0,0427
Time per week - use Internet on tablet	0,0422	-0,0012	0,0205	0,0271	0,0398
Bowling - Regularly do	0,0296	-0,0088	0,0205	0,0239	0,0041
Am seen as a risk taker	0,0927	-0,0072	0,0203	-0,0468	-0,0124
Rugby - Regularly do	0,027	0,003	0,02	0,0509	-0,0045
I like having lots to do	0,0667	0,0463	0,02	-0,0605	-0,0029
Will buy Blu-ray DVD player	0,0371	-0,0179	0,0199	0,0612	-0,0043
Films - Sci-Fi	-0,0113	0,0557	0,0194	0,0838	-0,0051
Will buy eReader	0,0312	-0,0103	0,0192	0,0527	-0,0109
CrossFit - Regularly do	0,0339	-0,0406	0,0188	0,065	-0,0107
Proud of roots/heritage	0,0038	0,108	0,0186	-0,0478	0,0186
Will buy streaming device	0,0363	0,0098	0,0182	0,0573	0,0176
Use other smartphone	0,0098	-0,0094	0,0181	0,0243	-0,019
Triathlons - Regularly do	0,0329	-0,0289	0,018	0,0511	0,002
How often do you do household shopping	0,0239	-0,0028	0,018	-0,0207	-0,0541
Travel by metro	0,0911	-0,0975	0,0178	0,053	-0,0795
Go to bars/pubs - Do regularly	0,0411	0,0279	0,0177	0,063	-0,0195
Swimming - Regularly do	0,0227	0,052	0,0173	0,0101	-0,0534
Mountain biking - Regularly do	0,0291	-0,0204	0,0168	0,0493	-0,0059
Will buy iPod or other music player	0,0379	-0,0153	0,0168	0,0477	-0,0042
Legal	0,0091	-0,005	0,0164	0,0034	0,0058
Equestrian excluding horse racing - Regularly do	0,03	-0,0121	0,016	0,0393	-0,0116
Will buy VR headset	0,0454	-0,0095	0,0159	0,0663	0,017
Ice Skating - Regularly do	0,026	-0,0089	0,0158	0,0424	-0,001
Use VR headset	0,0335	-0,0064	0,0156	0,0742	0,0309
Used online payment system recently	0,0097	0,0626	0,0156	0,0397	0,0286
Spend time with friends - Do regularly	-0,0004	0,107	0,0155	0,0347	-0,038
Films - Adventure	-0,0066	0,0747	0,0153	0,0649	-0,0016
Motivations behind newspaper reading - to escape from reality	0,0437	-0,0299	0,0145	0,0419	-0,0112
I prefer indie films/movies to mainstream movies	0,0857	-0,0532	0,0145	-0,0132	-0,0831
Gambling - Do regularly	0,0377	-0,0049	0,0143	0,0567	0,0165
People ask my advice before purchases	0,0897	0,0406	0,0143	-0,0392	0,0364
Relax at home - Do regularly	-0,0438	0,1194	0,0142	0,035	-0,0166
Armed Forces/Emergency Services/Security	0,0051	-0,0036	0,0142	0,0033	0,012
Motivations behind radio listening - to be inspired	0,0376	0,0241	0,014	0,0429	-0,0382
Cinema - visits per month	0,0612	-0,0094	0,0139	0,059	-0,0164
Pharmacy - visits per month	0,014	0,0332	0,0135	0,019	-0,0114

Boxing - Regularly do	0,0372	-0,0189	0,0133	0,0478	-0,0081
Motivations behind magazine reading - to escape from reality	0,0469	-0,0113	0,0133	0,0271	-0,0196
Have VR headset	0,0329	-0,0055	0,0131	0,0686	0,0437
Will buy older gaming handheld	0,0488	-0,026	0,013	0,0548	0,004
Used social media payment recently	0,0315	0,0365	0,013	0,0333	0,0448
Transportation/Logistics	-0,0042	-0,0103	0,0128	0,008	0,012
Travel by train	0,0863	-0,0784	0,0127	0,062	-0,0865
Use ZTE smartphone	0,0182	-0,0328	0,0126	0,048	-0,0244
Business/Finance Management	0,013	-0,0187	0,0125	-0,0066	0,0209
Province	0,0008	-0,0133	0,0125	-0,0092	-0,0396
Will buy wearable fitness device	0,0288	0,0277	0,0123	0,043	0,0209
Motivations behind TV watching - to be entertained	-0,0465	0,1125	0,0123	0,0358	0,0023
Use Huawei smartphone	0,0234	-0,0364	0,0123	0,0341	-0,0096
Prefer local shops to big supermarket	0,0557	0,0032	0,0123	-0,0515	-0,0653
Use Windows phone	0,0238	-0,0388	0,0118	0,032	-0,0017
Seize every opportunity	0,0458	0,0909	0,0117	-0,0715	-0,0507
Bars/Pubs/Clubs - visits per month	0,0487	0,0041	0,0115	0,0488	-0,0145
Retired	-0,0039	-0,001	0,0115	-0,0024	-0,0187
Time per week - walking	0,0241	0,0189	0,0112	0,0513	-0,0644
I shop around for value for money	0,0136	0,0706	0,0109	-0,065	-0,0421
Time per week - watch TV	-0,0197	0,0469	0,0108	0,0305	0,0135
Time per week - travel by metro	0,0682	-0,0733	0,0105	0,0675	-0,0708
Use Sony smartphone	0,0328	-0,0428	0,0103	0,0355	-0,0223
I entertain a lot	0,1011	0,008	0,0103	-0,0416	0,0256
Very ambitious	0,0889	0,0489	0,0101	-0,082	-0,0034
I am willing to spend money to save time	0,0695	0,0185	0,01	-0,0287	0,0763
American football - Regularly do	0,0339	0,0078	0,0099	0,044	-0,0056
Education/Training	-0,009	0,0091	0,0097	-0,0093	-0,0342
Motivations behind cinema trips - to be entertained	-0,0328	0,1043	0,0095	0,0503	-0,0075
Rock/Pop concert venues - visits per month	0,063	-0,0455	0,0086	0,053	-0,024
Films - Action	-0,0137	0,0676	0,0085	0,0403	0,0422
Will buy smartphone	0,0282	0,0173	0,0083	0,0352	0,0408
Have smartphone	-0,0553	0,1258	0,0083	0,0162	0,0324
Religion	0,0784	-0,0454	0,0082	-0,0485	-0,0161
Will buy laptop/desktop	0,0253	0,0105	0,0079	0,036	0,0027
Use Samsung smartphone	-0,0002	-0,0046	0,0077	0,0252	-0,0112
Agriculture	0,0049	-0,0178	0,0073	-0,0003	-0,003
I spend more for good food	0,0512	0,0883	0,007	-0,0521	-0,0237
Use Google Pixel smartphone	0,033	-0,0372	0,0069	0,0322	0,0042
I always look for information about a new film before seeing it	0,0442	0,0628	0,0069	-0,0311	-0,0427

Enjoy life as it happens	0,0533	0,0191	0,0068	-0,041	-0,0481
Career is very important	0,0623	0,0592	0,0067	-0,0477	0,0347
Organic products are healthier	0,0709	0,015	0,0067	-0,0508	-0,0579
Motivations behind internet use - to be informed	-0,031	0,1107	0,0065	0,0303	0,0028
Important to be an individual	0,0261	0,0793	0,0062	-0,0424	-0,0772
Used debit card recently	0,037	0,0021	0,006	0,0403	0,0023
Time per week - travel by train	0,0662	-0,0598	0,0058	0,0779	-0,0729
Used virtual currency recently	0,0515	-0,0432	0,0058	0,0399	-0,0019
Used own bank's own website/app to pay recently	0,0103	0,0634	0,0057	0,0491	0,0404
Use HTC smartphone	0,0269	-0,0335	0,0052	0,0254	-0,0183
Used prepaid card recently	0,0492	-0,0369	0,0049	0,0439	0
Motivations behind internet use - other reason	0,0085	-0,0109	0,004	0,0158	-0,0272
Creative	0,0658	0,0608	0,004	-0,0493	-0,0883
Motivations behind cinema trips - to be informed	0,0556	-0,0203	0,0033	0,0422	-0,0171
Will buy older gaming console	0,0341	-0,0301	0,0032	0,058	-0,0064
Important to follow tradition/customs	0,0659	0,0047	0,0032	-0,0767	0,0023
Motivations behind newspaper reading - connect with friends and family	0,0529	-0,0412	0,003	0,0334	-0,0092
Shopping - Do regularly	0,0087	0,0913	0,0029	0,0207	-0,0006
Motivations behind social media use - other reason	0,0081	-0,0221	0,0028	0,0168	-0,0128
Motivations behind TV watching - to relax/unwind	-0,0384	0,1064	0,0024	0,0438	0,0134
Motivations behind cinema trips - to escape from reality	-0,0016	0,0863	0,0023	0,0638	-0,0196
Important that I enjoy what I eat	0,0126	0,1169	0,0023	-0,06	-0,0489
Make an active effort to make new friends	0,1036	0,0081	0,0022	-0,0624	-0,0298
Industry	-0,004	0,0037	0,002	0,02	0,0321
Motivations behind magazine reading - connect with friends and family	0,0597	-0,0445	0,0017	0,0338	-0,0056
Motivations behind newspaper reading - to fill spare time	0,0171	0,0304	0,0014	0,0574	-0,006
Use Apple iPhone	0,0188	0,0383	0,001	-0,0248	0,0649
Films - Thriller	0,0009	0,0672	0,0005	0,0664	-0,0175
Motivations behind cinema trips - to relax/unwind	-0,0002	0,0883	-0,0004	0,0586	-0,0103
Automotive	0,0064	-0,0161	-0,0009	-0,0051	0,01
Motivations behind radio listening - connect with friends and family	0,0528	-0,0345	-0,0011	0,0439	-0,006
Motivations behind radio listening - to escape from reality	0,0249	0,0287	-0,0013	0,0344	-0,0151
Motivations behind cinema trips - to be inspired	0,0381	0,0258	-0,0016	0,0593	-0,0448
Crafts - Do regularly	0,0263	0,0395	-0,0024	0,0438	-0,0529
Open to sharing economy services	0,0701	0,058	-0,0026	-0,0527	-0,0916
Films - Dark Comedy	0,0157	0,0563	-0,0028	0,0958	-0,0316
Tech is important for productivity	0,073	0,0637	-0,0033	-0,0218	0,065
Listen to music at home - Do regularly	0,004	0,0988	-0,0037	0,048	-0,0477
Prefer factual advertising	0,0275	0,0874	-0,0037	-0,0381	-0,0072

Important to think of new ideas and be creative	0,0724	0,0639	-0,004	-0,0479	-0,0877
Motivations behind cinema trips - do not do this activity	-0,0285	-0,018	-0,0042	-0,0196	-0,0182
Telecommunications	0,006	-0,0037	-0,0049	0,0021	0,0237
Buy clothes for comfort not style	0,0048	0,0079	-0,0053	-0,0144	-0,0233
I do not know what ad blocking software is	-0,0256	0,0014	-0,0054	-0,0239	-0,0129
Motivations behind TV watching - do not do this activity	-0,0051	-0,013	-0,0055	-0,0056	-0,0356
Important to get married	0,0691	-0,0127	-0,0057	-0,0601	0,0443
Motivations behind TV watching - to be inspired	0,0437	0,0072	-0,0058	0,0398	-0,0016
Films – none of the above	-0,0047	-0,0233	-0,0059	-0,0113	0,0003
Motivations behind magazine reading - to fill spare time	0,0076	0,0616	-0,0061	0,0491	0,0017
I save up for things I want	0,029	0,0484	-0,0072	-0,0488	-0,0558
Community/Social Services	-0,0004	0,0031	-0,0074	-0,0055	-0,0249
Healthcare/Medical/Pharmaceutical	-0,0125	0,0249	-0,0079	-0,023	-0,0021
Time spent on bus last time used	0,0646	-0,0534	-0,0083	0,0828	-0,0961
Films - 3D	0,042	0,0161	-0,0083	0,0593	0,0078
Motivations behind radio listening - to fill spare time	0,0074	0,053	-0,0085	0,0496	0,0239
Important to take risks and go on adventures	0,0967	0,0236	-0,0085	-0,0472	-0,0429
Personal appearance says a lot	0,0772	0,0405	-0,0085	-0,0674	0,0051
Consider myself very stylish	0,1075	0,0151	-0,0094	-0,0735	-0,0044
I budget carefully when shopping	0,0377	0,0139	-0,0095	-0,0566	-0,077
Consider myself expert in new tech	0,1028	-0,0164	-0,0096	-0,001	0,0961
Marketing/PR/Advertising	0,0093	0,0017	-0,0098	0,002	-0,0213
I plan for others often	0,0855	0,0254	-0,0098	-0,0553	-0,0022
Arts/Entertainment/Sports/Journalism/Media	-0,0004	0,0103	-0,0099	0,02	-0,052
Consider all options before purchase	0,0334	0,0555	-0,01	-0,0444	-0,0361
Follow a well-organised routine	0,0427	0,0374	-0,0101	-0,0565	0,027
Will buy none of these	-0,0762	0,013	-0,011	-0,0498	-0,0464
Films - Musical	0,0114	0,0501	-0,0113	0,0446	-0,0512
Motivations behind radio listening - other reason	0,0013	-0,0158	-0,0113	0,0174	-0,0101
Rather buy on credit than wait	0,0136	0,0775	-0,0114	-0,0552	-0,046
Motivations behind TV watching - connect with friends and family	0,0434	-0,0001	-0,0117	0,0454	0,0012
Motivations behind TV watching - other reason	0,0192	-0,0293	-0,0122	0,0175	-0,0042
None of these - Do regularly	0,003	-0,0411	-0,0135	-0,0101	0,0056
How close to seeing a film at the cinema did you decide what film to see?	-0,0199	0,0034	-0,0138	-0,0084	0,0109
Own brands are as good as branded	0,0191	0,045	-0,0141	-0,0347	-0,0453
Will buy latest gaming console	0,0292	0,0183	-0,0149	0,0776	0,0263
Early trend knowledge	0,1174	-0,0088	-0,0149	-0,0414	0,0476
Motivations behind cinema trips - other reason	0,0187	-0,0229	-0,0155	0,0212	-0,0053
Shopping centres - visits per month	0,0249	0,0432	-0,0156	0,0277	0,0087

Films - Children's Movies	0,0207	0,0214	-0,016	0,0425	0,0141
Only buy well-known brands	0,0866	-0,0291	-0,0163	-0,0495	0,0679
Have none of these	-0,002	-0,0255	-0,0165	-0,0246	-0,0117
Use older gaming handheld	0,0248	0,0097	-0,0168	0,0916	0,0155
Motivations behind TV watching - to escape from reality	-0,0131	0,0723	-0,0169	0,0541	0,0068
Films - Supernatural Thriller	0,0203	0,0467	-0,0177	0,0832	-0,0285
Films - Comedy	-0,0147	0,0905	-0,0182	0,0366	0,018
Use LG smartphone	0,0026	-0,0143	-0,0187	0,0207	-0,0339
I am passionate about films	0,0783	0,0326	-0,0189	0,0037	-0,0421
Motivations behind cinema trips - to fill spare time	0,0476	-0,0033	-0,0194	0,0548	-0,0022
Live alone	-0,0092	-0,0352	-0,0194	0,0155	-0,0961
Mix cheap and expensive clothing	0,0603	0,0197	-0,0202	-0,0267	0,0297
Motivations behind internet use - connect with friends and family	0,0131	0,0623	-0,0208	0,038	-0,0119
Live with grandparents	0,0084	-0,0033	-0,0211	0,0067	0,0003
Films - Fantasy	0,008	0,0465	-0,0212	0,0938	-0,0192
I treat myself to luxury items	0,0988	-0,0085	-0,0215	-0,0309	0,084
I don't expect to be much wealthier in future	0,0372	-0,0622	-0,0217	-0,0092	-0,007
Use latest gaming handheld	0,0301	-0,0014	-0,0221	0,0969	0,0405
Have latest gaming handheld	0,0232	0,0139	-0,0235	0,0879	0,0573
Hospitality/Leisure/Tourism	0,0008	0,0079	-0,0235	0,0044	-0,0176
I notice what is advertised on posters/digital screens	0,0924	0,0277	-0,0236	-0,0121	0,0013
Have older gaming console	-0,0221	0,0796	-0,0249	0,0821	0,0807
Have older gaming handheld	0,0112	0,0346	-0,0252	0,0883	0,0533
Never have enough free time	0,0424	0,0353	-0,0254	-0,0399	0,021
Like to share everyday adventures	0,0666	0,0724	-0,0254	-0,0558	-0,0295
Diet frequently to lose weight	0,074	-0,0218	-0,0258	-0,0408	0,0406
Love new gadgets/appliances	0,097	0,0285	-0,0267	0,0076	0,1166
Films - Romantic Comedies	-0,0055	0,0888	-0,0269	0,0301	-0,0195
Prefer to try/use free sample	0,0731	0,024	-0,0275	-0,0542	-0,0061
Have latest gaming console	0,0123	0,0513	-0,0276	0,0784	0,098
I'd rather wait or a film to go out on TV/DVD/streaming than go to the cinema	0,0239	-0,0336	-0,0284	-0,043	0,0115
Believe in traditional family roles	0,0828	-0,0449	-0,0285	-0,0626	0,0295
Use older gaming console	-0,002	0,0391	-0,0289	0,1025	0,042
I often look to see if posters/digital screens on my usual routes have changes	0,1086	-0,0152	-0,0289	-0,0171	0,0176
Films - Animation	0,0216	0,0323	-0,029	0,0806	-0,0126
Keep my feelings to myself	0,0235	-0,0161	-0,029	0,002	0,0255
Buy newest brands/styles fashion	0,1241	-0,0118	-0,0296	-0,0621	0,0518
Time per week - travel by bus	0,0515	-0,0404	-0,0298	0,0786	-0,0705
Like to try new brands	0,0833	0,0507	-0,0302	-0,0561	0,0044

Gaming - Do regularly	0,0201	0,0345	-0,0304	0,1097	0,0344
I want to be popular	0,095	-0,0053	-0,0306	-0,0413	0,0544
Find it hard to fit everything into my day	0,0493	0,0419	-0,0307	-0,0273	0,0109
Advertising helps me	0,0519	0,068	-0,0308	-0,0452	0,0318
Fast food restaurants - Do regularly	0,0091	0,0562	-0,0313	0,0738	0,019
Motivations behind TV watching - to fill spare time	-0,0105	0,0703	-0,0314	0,0647	0,0277
Internet activities	0,0254	0,0268	-0,0315	0,0115	0,0728
Motivations behind magazine reading - other reason	-0,0015	-0,021	-0,0316	0,0137	0,0034
Expect advertising to be entertaining	0,0569	0,0572	-0,0324	-0,0464	0,0212
Posters/digital screens make my journey more interesting	0,099	0,0094	-0,0325	-0,0161	0,0132
Wear sportswear when not exercising	0,0682	-0,0016	-0,0326	-0,0189	0,0712
Films - Romance	0,0105	0,0724	-0,0348	0,0169	-0,0198
Important to be successful; to have people recognize them	0,0904	0,0308	-0,036	-0,0675	0,0286
Prefer brands that reward loyalty	0,0398	0,0824	-0,0363	-0,0571	0,0094
Motivations behind social media use - connect with friends and family	-0,0116	0,1082	-0,0365	0,0386	-0,0025
Motivations behind newspaper reading - other reason	-0,0011	-0,0208	-0,0368	0,0118	0,0117
Travel by bus	0,0777	-0,0643	-0,0376	0,0764	-0,1071
Motivations behind internet use - to be inspired	0,0151	0,0833	-0,0376	0,0462	-0,0206
Motivations behind cinema trips - connect with friends and family	0,0243	0,0479	-0,038	0,0413	-0,0134
Friends most important thing in my life	0,0794	0,0028	-0,038	-0,041	-0,0491
Use latest gaming console	0,0218	0,0262	-0,0395	0,0915	0,0719
Important to have a good time/spoil him or herself	0,0662	0,0348	-0,0398	-0,0364	-0,0021
Stick to brands I Like	0,0311	0,0543	-0,0398	-0,0409	0,0283
Internet activities	0,056	-0,0139	-0,0424	0,0848	0,0607
Important to be reach, to have a lot of money	0,1043	-0,0384	-0,0427	-0,0334	0,0826
Look to my friends for inspiration	0,0909	0,0398	-0,0429	-0,0481	-0,0245
Others' opinions are important to me	0,0705	0,0548	-0,0442	-0,0487	0,0168
Prefer fewer but higher quality purchases	0,0526	0,0578	-0,0464	-0,0358	-0,0275
I spend a lot on clothes for myself	0,1112	-0,0312	-0,0477	-0,0332	0,0547
International brands have status	0,092	0,0058	-0,0488	-0,0407	0,0336
I regularly use mobile coupons	0,0988	-0,0017	-0,049	-0,0198	0,0608
Motivations behind internet use - to be entertained	-0,0149	0,0978	-0,0499	0,0625	0,0083
Food Prep/Service	0,0034	-0,0042	-0,0505	0,0113	-0,0283
Influenced by deals	0,0401	0,07	-0,0507	-0,0481	0,0024
Motivations behind internet use - to relax/unwind	-0,0018	0,0823	-0,0512	0,0716	0,0159
I like to have the latest mobile phone	0,1071	0,0073	-0,0515	-0,0067	0,131
Money is best measure of success	0,0917	-0,0611	-0,0519	-0,0309	0,0822
Time per week - play games on PC/console	0,0353	-0,025	-0,0536	0,0922	0,0295
Films - Horror	0,0175	0,0381	-0,0537	0,0688	-0,0117

Fatalistic	0,0662	-0,0862	-0,054	-0,014	0,038
Motivations behind social media use - to be informed	0,0241	0,067	-0,0541	0,0504	0,0092
Spend a lot on beauty/grooming	0,105	-0,0016	-0,0545	-0,0364	0,0365
Motivations behind internet use - to fill spare time	-0,0166	0,0841	-0,0549	0,0632	0,0265
Motivations behind magazine reading - do not do this activity	-0,0357	-0,0274	-0,0561	-0,0097	0,0098
Live with other adults	0,0079	-0,0003	-0,0583	0,0336	-0,0544
Sales/Retail	-0,0018	0,0179	-0,0584	0,0189	-0,0129
Internet activities	0,0312	0,0557	-0,0584	0,0166	0,1
Enjoy browsing in shops	0,0609	0,0646	-0,0588	-0,0343	-0,0036
Live ONLY with other adults	0,0079	-0,0008	-0,0591	0,0277	-0,0636
Motivations behind radio listening - do not do this activity	-0,0086	-0,0285	-0,0596	0,0098	-0,0478
Motivations behind social media use - to be inspired	0,0369	0,0621	-0,06	0,0464	-0,0073
None of these - Regularly do	-0,0504	-0,0229	-0,0605	0,01	0,0339
Motivations behind newspaper reading - do not do this activity	-0,0205	-0,0207	-0,0618	-0,0108	0,0043
Motivations behind internet use - to escape from reality	0,0101	0,0618	-0,0637	0,0747	0,0146
Trust professional brand endorsements	0,0917	0,0095	-0,0637	-0,05	0,0295
I ask others for opinions more than they ask me	0,074	-0,0347	-0,0643	-0,0225	0,0121
More likely to buy if celebrity I like endorses	0,1073	-0,0419	-0,0688	-0,0353	0,0529
Prefer others' approval of brands I buy	0,0949	-0,0122	-0,0695	-0,0413	0,0646
Motivations behind social media use - to escape from reality	0,0238	0,0479	-0,0712	0,0624	0,0085
Time per week - online videos	0,0528	-0,0081	-0,0714	0,0792	-0,0098
Influenced by others' opinions	0,0711	0,0314	-0,0722	-0,0432	0,0423
Time per week - watch TV/video streamed on demand	0,0298	0,0487	-0,0735	0,0746	0,0206
Early adopter	0,0436	0,0418	-0,0739	-0,0193	-0,005
Internet activities	0,0438	-0,0048	-0,0758	0,1043	0,0513
Motivations behind social media use - to relax/unwind	0,014	0,0792	-0,0763	0,0597	0,017
Motivations behind social media use - to be entertained	0,0055	0,0971	-0,078	0,0518	0,0179
Not comfortable to meet new people	0,0213	-0,0404	-0,0788	0,019	0,0324
I spend money without thinking	0,0769	-0,0282	-0,0788	0,0122	0,0862
My mobile phone is my main source of entertainment	0,0773	0,0333	-0,0815	-0,0102	0,1078
Concerned what others think	0,0728	-0,0039	-0,0848	-0,0337	0,0472
Internet activities	0,0333	0,056	-0,0881	0,0209	0,0334
Prefer natural skincare products	0,052	0,0192	-0,0882	-0,026	0,0265
Motivations behind social media use - to fill spare time	0,0004	0,0953	-0,0883	0,0642	0,021
Attend school/further education	0,0051	0,0161	-0,0897	0,0286	-0,039
Work part time	-0,0061	0,0092	-0,0923	0,0269	-0,0994
Do not rent or own	-0,0018	0,0116	-0,0931	0,0289	-0,0087
I feel lost without my smartphone	0,0685	0,0153	-0,0942	0,0426	0,0838
Live with parents	0,0186	0,0066	-0,0963	0,0281	0,0012
Time per week - smartphone	0,0429	0,0379	-0,0964	0,0603	0,0437

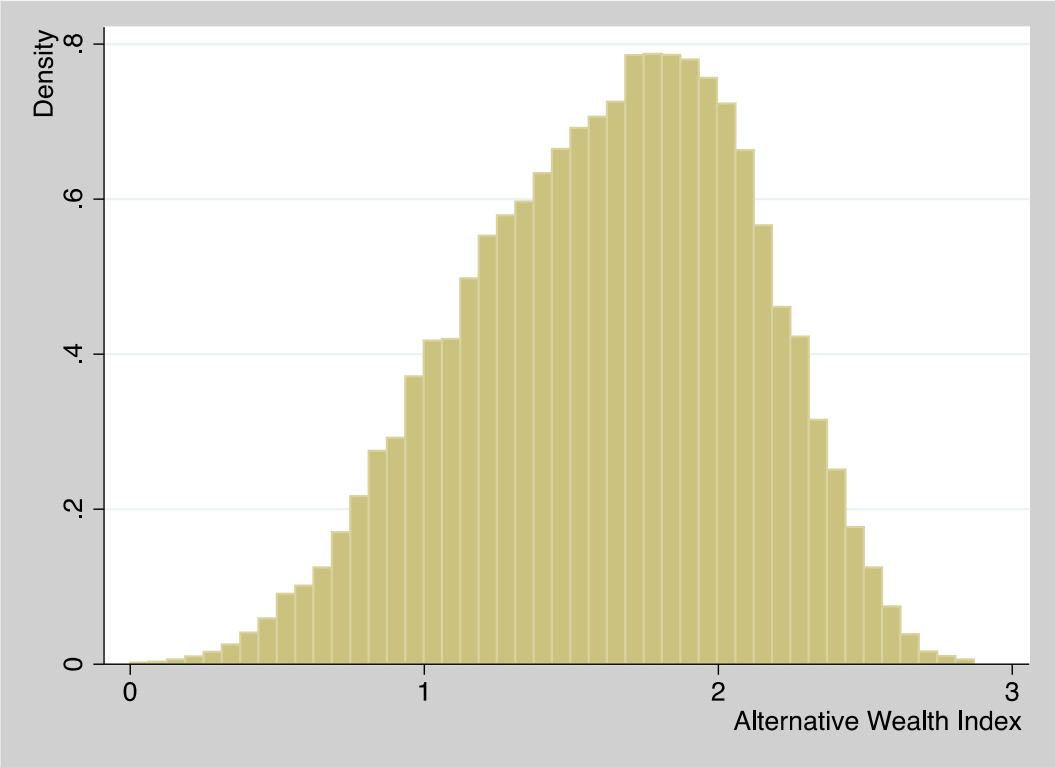
Time per week - social media	0,038	0,0348	-0,0991	0,0567	0,0077
Viewed as worrier	0,0559	-0,0214	-0,0996	-0,0051	0,015
Have not read newspapers in last year	-0,0348	-0,007	-0,0998	-0,0262	0,0259
Rent house/apartment	0,0126	-0,0007	-0,1055	0,048	-0,1189
My mobile phone is an important part of my social life	0,09	0,0511	-0,106	-0,0071	0,0861
I worry a lot about money	**0,0331**	**0,0121**	**-0,1112**	**0,0038**	**-0,001**
Mobile is my main source of entertainment	0,0931	0,0138	-0,1204	-0,0035	0,1016

Appendix 3: Alternative Wealth Index ('shadow analysis')

3.1 Construction formula for the Alternative Wealth Index:

$$AWI_{i,t} = 0.1788ownhome_{i,t} + 0.1487fulltime_{i,t} + 0.1323education_{i,t} + 0.0864exercise_{i,t} + 0.758balancedmeals_{i,t} - 0.0972donotreadnewspaper_{i,t} - 0.0851smartphoneuse_{i,t} - 0.0814worryaboutmoney_{i,t}$$

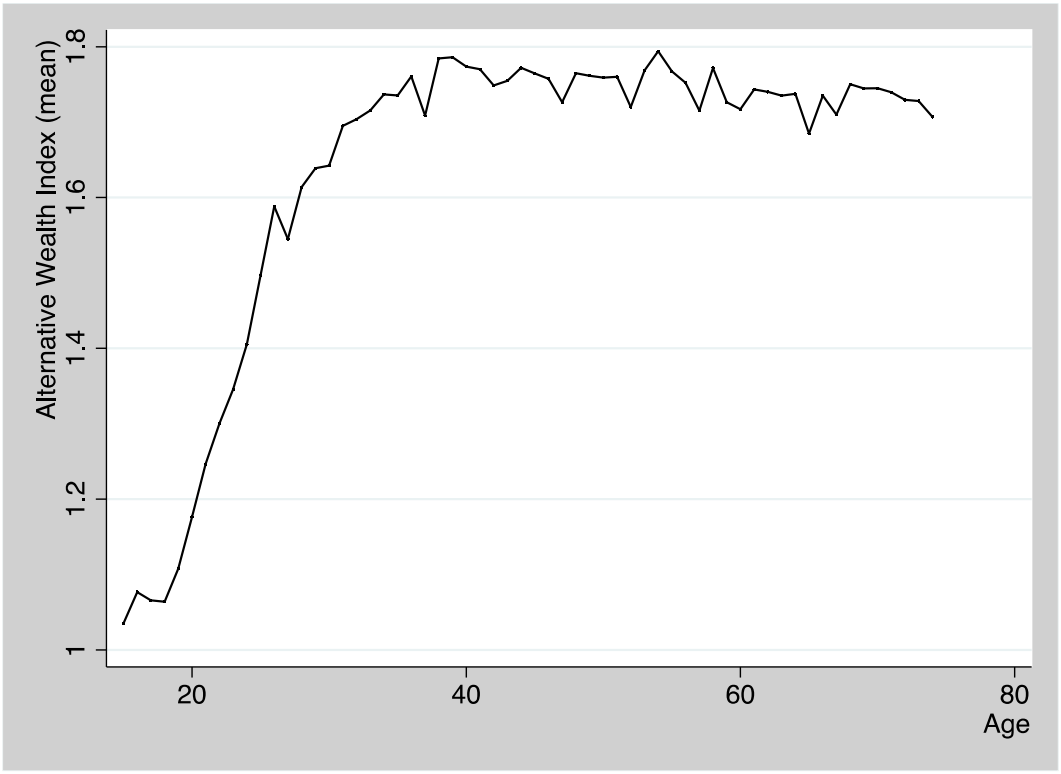
3.2 Alternative Wealth Index distribution

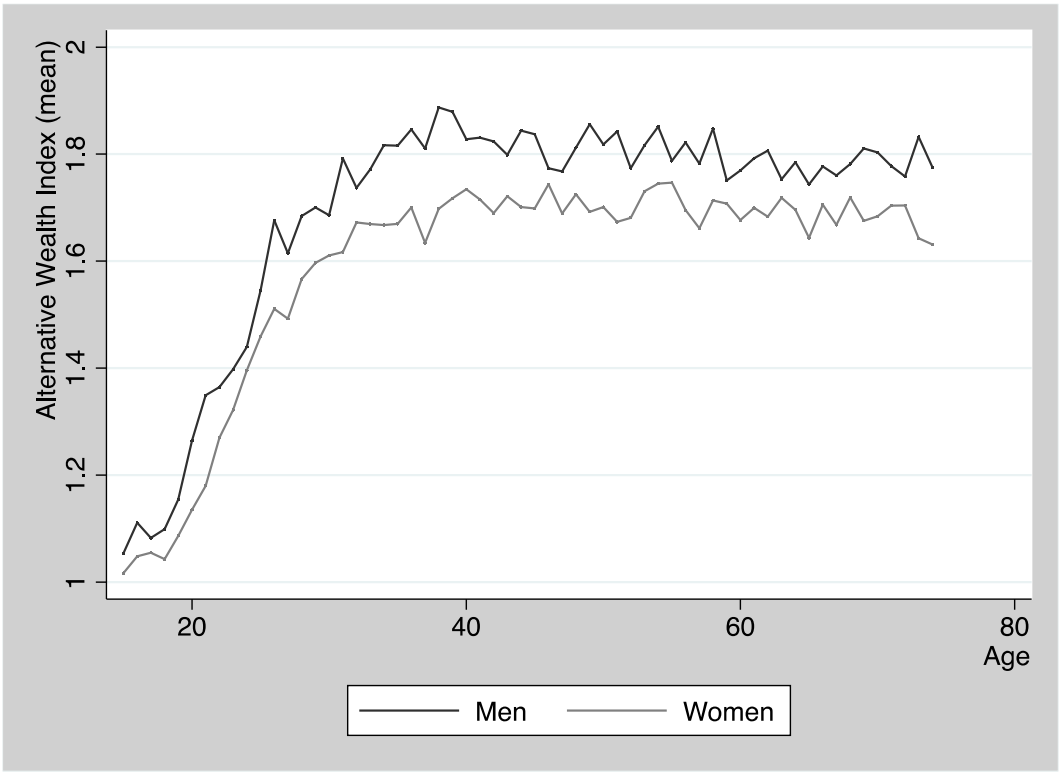


3.3 Descriptive statistics of the Alternative Wealth Index

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max	(6) p25	(7) median	(8) p75
WI	40,326	1.622	0.479	0	2.869	1.283	1.66	1.982

3.4 Selected age profiles with the Alternative Wealth Index





3.5 First stage IV regression results

Note: education is not included as a covariate because it is a variable in the index construction formula of the Alternative Wealth Index

	(1)	(2)
VARIABLES	P1	P2
Single dummy	0.0154	0.0443***
	(0.0104)	(0.0112)
Age	0.00768***	0.0190***
	(0.00181)	(0.00194)
Square of age ÷100	-0.005.61***	-0.0163***
	(0.00205)	(0.00219)
White dummy	-0.0669***	-0.0500***
	(0.0124)	(0.0132)
Rural dummy	-0.00257	0.0476***
	(0.0100)	(0.0107)
Routine	0.158***	0.191***
	(0.00487)	(0.00521)
Plan for friends	0.0714***	0.0724***
	(0.00461)	(0.00493)
Inverse of 'No plan'	-0.0426***	-0.0786***
	(0.00472)	(0.00504)
Constant	2.848***	2.354***
	(0.0463)	(0.0495)
Observations	40,239	40,239
R-squared	0.039	0.052

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

3.6 Second stage IV regression results

VARIABLES	(1) P1	(2) P2
Planning (P)	0.317***	0.266***
	(0.0145)	(0.0119)
Single dummy	-0.406***	-0.413***
	(0.00519)	(0.00507)
Age	0.0348***	0.0323***
	(0.000912)	(0.000912)
Square of age ÷100	-0.0336***	-0.0312***
	(0.00103)	(0.00102)
Male dummy	0.110***	0.125***
	(0.00484)	(0.00483)
White dummy	0.0315***	0.0252***
	(0.00626)	(0.00606)
Rural dummy	-0.0611***	-0.0748***
	(0.00501)	(0.00488)
Constant	-0.208***	0.0860**
	(0.0554)	(0.0420)
Observations	39,784	39,784
R-squared	0.009	0.056

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Implied percentage impact of a one-point increase planning on the Alternative Wealth Index:

- Using P1: 19.6%
- Using P2: 16.4%

3.7 Regression results using year dummy interaction terms

VARIABLES	(1) HWI
P2017	0.289***
	(0.0141)
P2021	0.332***
	(0.0142)
Single dummy	-0.413***
	(0.00507)
Age	0.0353***
	(0.000890)
Square of age ÷100	-0.0341***
	(0.0010)
Male dummy	0.100***
	(0.00474)
White dummy	0.0442***
	(0.00613)
Rural dummy	-0.0567***
	(0.00490)
Constant	-0.190***
	(0.0541)
Observations	39,784
R-squared	0.054

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Implied percentage impact of a one-point increase planning on the Alternative Wealth Index:

- In 2017: 18.4%
- In 2021: 19.7%
- Increase in relationship strength: 7.1%

3.8 Change in mean and median Alternative Wealth Index, by propensity to plan

Self-reported propensity to plan	Mean Alt. Wealth Index, 2017 sample	Mean Alt. Wealth Index, 2021 sample	% change
5 (highest)	1.582	1.696	+7.2%
4	1.600	1.730	+8.1%
3	1.530	1.645	+7.5%
2	1.528	1.602	+4.8%
1 (lowest)	1.471	1.516	+3.1%
All respondents	*1.571*	*1.684*	*+7.1%*

Self-reported propensity to plan	Median Alt. Wealth Index, 2017 sample	Median Alt. Wealth Index, 2021 sample	% change
5 (highest)	1.629	1.725	+5.9%
4	1.651	1.780	+7.8%
3	1.575	1.682	+6.8%
2	1.570	1.599	+1.8%
1 (lowest)	1.513	1.517	+0.3%
All respondents	*1.617*	*1.722*	*+6.5%*

3.9 Quantile regression results

Note: Please see Appendix 4 for core Wealth Index full results

VARIABLES	(1) 10%	(2) 20%	(3) 30%	(4) 40%	(5) 50%	(6) 60%	(7) 70%	(8) 80%	(9) 90%
P	0.415***	0.391***	0.344***	0.327***	0.335***	0.305***	0.280***	0.267***	0.246***
	(0.0213)	(0.0212)	(0.0192)	(0.0177)	(0.0157)	(0.0165)	(0.0196)	(0.0215)	(0.0426)
Single dummy	-0.434***	-0.430***	-0.424***	-0.418***	-0.414***	-0.407***	-0.401***	-0.391***	-0.379***
	(0.0116)	(0.00812)	(0.0181)	(0.00728)	(0.00729)	(0.00646)	(0.00855)	(0.00672)	(0.00831)
Education	0.0279	0.0306***	0.0331	0.0361	0.0376	0.0389	0.0386	0.0382***	0.0366***
	(0.0301)	(0.00938)	(2.592)	(1.230)	(0.591)	(0.558)	(1.123)	(0.00552)	(0.0127)
Age	-0.000240	-0.000277	-0.000308	-0.000347	-0.000368	-0.000386	-0.000386	-0.000384***	-0.000374**
	(0.00144)	(0.000308)	(0.0426)	(0.0173)	(0.00727)	(0.00593)	(0.0106)	(8.27e-05)	(0.000183)
Age squared	0.141***	0.129***	0.122***	0.117***	0.113***	0.107***	0.0999***	0.0883***	0.0757***
	(0.00991)	(0.00874)	(0.00684)	(0.00620)	(0.00618)	(0.00600)	(0.00585)	(0.00616)	(0.00751)
Male dummy	0.0285***	0.0375***	0.0328***	0.0338***	0.0373***	0.0332***	0.0296***	0.0322**	0.0263
	(0.00834)	(0.00811)	(0.0105)	(0.00638)	(0.00797)	(0.00822)	(0.0101)	(0.0152)	(0.0223)
White dummy	-0.0593***	-0.0686***	-0.0652***	-0.0662***	-0.0629***	-0.0619***	-0.0583***	-0.0635***	-0.0620***
	(0.0104)	(0.00852)	(0.0113)	(0.00666)	(0.00649)	(0.00610)	(0.00634)	(0.00630)	(0.00756)
Rural dummy	-1.162	-0.855	-0.536	-0.389	-0.324	-0.109	0.113	0.311	0.618
	(0)	(0)	(39.20)	(21.77)	(11.96)	(13.06)	(29.60)	(0)	(0)
Observations	39,784	39,784	39,784	39,784	39,784	39,784	39,784	39,784	39,784

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix 4: Full quantile regression Results

4.1 Dependent variable: Wealth Index

VARIABLES	(1) 10%	(2) 20%	(3) 30%	(4) 40%	(5) 50%	(6) 60%	(7) 70%	(8) 80%	(9) 90%
P	0.127***	0.109***	0.134***	0.143***	0.128***	0.125***	0.148***	0.152***	0.141
	(0.0421)	(0.0284)	(0.0454)	(0.0317)	(0.0299)	(0.0301)	(0.0342)	(0.0380)	(0.0948)
Single dummy	-0.251***	-0.246***	-0.279***	-0.280***	-0.265***	-0.234***	-0.204***	-0.177***	-0.145***
	(0.0373)	(0.0296)	(0.0301)	(0.0127)	(0.0134)	(0.0119)	(0.0150)	(0.0120)	(0.0143)
Education	0.112***	0.109***	0.147***	0.174***	0.191***	0.196***	0.187***	0.175***	0.155***
	(0.0205)	(0.0182)	(0.00460)	(0.00393)	(0.00412)	(0.00346)	(0.00328)	(0.00314)	(0.00377)
Age	0.0275	0.0225	0.0410	0.0541	0.0617	0.0708	0.0789	0.0823***	0.0786***
	(0.0744)	(4.896)	(4.602)	(3.177)	(1.406)	(1.069)	(1.718)	(0.00882)	(0.0202)
Square of age ÷100	-0.0258	-0.0201	-0.0401	-0.0542	-0.0618	-0.0710	-0.0781	-0.0805***	-0.0764***
	(0.415)	(0.952)	(0.770)	(0.452)	(0.175)	(0.115)	(0.164)	(0.0128)	(0.0271)
Male dummy	0.350***	0.356***	0.468***	0.513***	0.494***	0.447***	0.401***	0.376***	0.365***
	(0.0244)	(0.0517)	(0.0121)	(0.0114)	(0.0115)	(0.0110)	(0.0108)	(0.0113)	(0.0134)
White dummy	0.0977***	0.0913**	0.120***	0.139***	0.139***	0.139***	0.123***	0.117***	0.0986**
	(0.0245)	(0.0442)	(0.0153)	(0.0135)	(0.0180)	(0.0158)	(0.0182)	(0.0274)	(0.0439)
Rural dummy	-0.109***	-0.0999***	-0.144***	-0.172***	-0.174***	-0.169***	-0.159***	-0.161***	-0.171***
	(0.0197)	(0.0113)	(0.0229)	(0.0126)	(0.0121)	(0.0119)	(0.0125)	(0.0113)	(0.0135)
Constant	-1.309	-0.618	-1.014	-1.200	-1.161	-1.143	-1.148	-0.968	-0.442
	(0)	(62.35)	(68.34)	(55.54)	(28.09)	(24.72)	(44.90)	(0)	(0)
Observations	35,803	35,803	35,803	35,803	35,803	35,803	35,803	35,803	35,803

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

4.2 Dependent variable: Log of Wealth Index

VARIABLES	(1) 10%	(2) 20%	(3) 30%	(4) 40%	(5) 50%	(6) 60%	(7) 70%	(8) 80%	(9) 90%
P	0.171***	0.134***	0.130***	0.102***	0.0955***	0.0767***	0.0826***	0.0757***	0.0608**
	(0.0333)	(0.0338)	(0.0289)	(0.0252)	(0.0193)	(0.0179)	(0.0188)	(0.0208)	(0.0303)
Single dummy	-0.324***	-0.309***	-0.260***	-0.212***	-0.177***	-0.144***	-0.116***	-0.0933***	-0.0574***
	(0.0194)	(0.0129)	(0.0311)	(0.00924)	(0.00878)	(0.00738)	(0.00776)	(0.00650)	(0.00585)
Education	0.124***	0.136***	0.137***	0.133***	0.126***	0.116***	0.103***	0.0876***	0.0613***
	(0.00860)	(0.00395)	(0.00337)	(0.00284)	(0.00249)	(0.00204)	(0.00191)	(0.00232)	(0.00147)
Age	0.0194	0.0332*	0.0447	0.0527	0.0550	0.0576	0.0575	0.0515	0.0329***
	(0.0897)	(0.0200)	(6.821)	(2.064)	(0.580)	(0.305)	(0.360)	(0.720)	(0.00413)
Square of age ÷100	-0.0144	-0.0305	-0.0433	-0.0521	-0.0546	-0.0573	-0.0569	-0.0506	-0.0320***
	(0.433)	(0.0727)	(0.112)	(0.291)	(0.711)	(0.324)	(0.342)	(0.623)	(0.00407)
Male dummy	0.373***	0.412***	0.394***	0.347***	0.286***	0.230***	0.194***	0.170***	0.140***
	(0.0170)	(0.0139)	(0.00994)	(0.00831)	(0.00741)	(0.00657)	(0.00582)	(0.00626)	(0.00535)
White dummy	0.0976***	0.113***	0.107***	0.108***	0.0938***	0.0802***	0.0666***	0.0539***	0.0413**
	(0.0145)	(0.0128)	(0.0166)	(0.00885)	(0.00973)	(0.00908)	(0.00998)	(0.0145)	(0.0167)
Rural dummy	-0.128***	-0.119***	-0.122***	-0.119***	-0.113***	-0.0981***	-0.0866***	-0.0767***	-0.0664***
	(0.0176)	(0.0133)	(0.0239)	(0.00887)	(0.00801)	(0.00678)	(0.00645)	(0.00587)	(0.00528)
Constant	-2.313	-2.158	-2.134	-1.988	-1.803	-1.589	-1.417	-1.065	-0.318
	(0)	(0)	(103.9)	(36.44)	(11.77)	(7.144)	(9.402)	(20.74)	(0)
Observations	35,718	35,718	35,718	35,718	35,718	35,718	35,718	35,718	35,718

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix 5: Additional Regression Results

5.1 First-stage OLS results with each candidate variable shown individually

Candidate z_1 (routine):

VARIABLES	(1) P1	(2) P2
Single dummy	0.0112	0.0266**
	(0.0106)	(0.0113)
Education	0.0186***	-0.0157***
	(0.00273)	(0.00292)
Age	0.00326*	0.0195***
	(0.00189)	(0.00202)
Square of age ÷100	-0.001.57	-0.0174***
	(0.00212)	(0.00227)
Male dummy	-0.0229**	-0.0801***
	(0.00979)	(0.0105)
White dummy	-0.0648***	-0.0720***
	(0.0126)	(0.0134)
Rural dummy	-0.00915	0.0189*
	(0.0103)	(0.0110)
Routine (z1)	0.164***	0.199***
	(0.00489)	(0.00523)
Constant	2.982***	2.496***
	(0.0428)	(0.0458)
Observations	39,784	39,784
R-squared	0.032	0.042

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Candidate z_2 (plan for friends):

VARIABLES	(1) P1	(2) P2
Single dummy	0.0116	0.0263**
	(0.0107)	(0.0115)
Education	0.0196***	-0.0142***
	(0.00276)	(0.00296)
Age	0.00731***	0.0243***
	(0.00190)	(0.00204)
Square of age ÷100	-0.00565***	-0.023***
	(0.00214)	(0.0023)
Male dummy	-0.0325***	-0.0917***
	(0.00988)	(0.0106)
White dummy	-0.0593***	-0.0658***
	(0.0127)	(0.0136)
Rural dummy	0.000640	0.0297***
	(0.0104)	(0.0111)
Plan for friends (z2)	0.0913***	0.104***
	(0.00465)	(0.00499)
Constant	3.196***	2.784***
	(0.0435)	(0.0467)
Observations	39,784	39,784
R-squared	0.014	0.018

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Candidate z_3 (Inverse of 'No plan'):

VARIABLES	(1) P1	(2) P2
Single dummy	0.00265	0.0167
	(0.0107)	(0.0115)
Education	0.0255***	-0.00672**
	(0.00276)	(0.00296)
Age	0.00526***	0.0220***
	(0.00191)	(0.00205)
Square of age ÷100	-0.00456**	-0.0212***
	(0.00215)	(0.00230)
Male dummy	-0.0373***	-0.1000***
	(0.00993)	(0.0106)
White dummy	-0.0655***	-0.0723***
	(0.0127)	(0.0136)
Rural dummy	-0.0137	0.0136
	(0.0104)	(0.0111)
Plan for friends (z2)	-0.0457***	-0.0790***
	(0.00481)	(0.00516)
Constant	3.650***	3.364***
	(0.0415)	(0.0445)
Observations	39,784	39,784
R-squared	0.007	0.013

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

5.2 IV regressions for 2017 and 2021 samples, using z_2 as an instrument

VARIABLES	(1) 2017 sample	(2) 2021 sample
Planning (P)	0.442***	0.634***
	(0.0656)	(0.0747)
Single dummy	-0.211***	-0.267***
	(0.0137)	(0.0170)
Education	0.159***	0.138***
	(0.00368)	(0.00491)
Age	0.0502***	0.0586***
	(0.00245)	(0.00322)
Square of age ÷100	-0.0480***	-0.0600***
	(0.00277)	(0.00358)
Male dummy	0.395***	0.451***
	(0.0127)	(0.0163)
White dummy	0.126***	0.188***
	(0.0177)	(0.0206)
Rural	-0.138***	-0.161***
	(0.0130)	(0.0172)
Constant	-1.993***	-2.705***
	(0.244)	(0.265)
Observations	19,420	16,383
R-squared	0.056	-0.184

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

5.3 IV regression with alternative specification, using z_2 as an instrument

VARIABLES	(1) Alternative specification
Planning (P)	0.544***
	(0.0500)
2021 year dummy	0.0609***
	(0.0102)
Single dummy	-0.238***
	(0.0108)
Education	0.149***
	(0.00300)
Age	0.0546***
	(0.00197)
Square of age ÷100	-0.0541***
	(0.00222)
Male dummy	0.423***
	(0.0102)
White dummy	0.158***
	(0.0135)
Rural	-0.150***
	(0.0106)
Constant	-2.417***
	(0.183)
Observations	35,803
R-squared	-0.068

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

5.4 First-stage IV results for the regression of news readership on the planning variable

Note: instrument highlighted in grey

VARIABLES	(1) P
Single dummy	0.00767
	(0.0107)
Education	0.0187***
	(0.00278)
Age	0.00423**
	(0.00191)
Square of age ÷100	-0.00410*
	(0.00214)
Male dummy	-0.0427***
	(0.00993)
White dummy	-0.0700***
	(0.0127)
Rural dummy	-0.00830
	(0.0104)
Read news dummy	0.146***
	(0.0107)
Constant	3.506***
	(0.0400)
Observations	39,784
R-squared	0.009

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

5.5 Second-stage IV results for the regression of news readership on financial planning

VARIABLES	(1) P
Read finance/money Section dummy	0.503***
	(0.0371)
Single dummy	0.0109
	(0.0108)
Education	0.00868***
	(0.00301)
Age	0.00605***
	(0.00193)
Square of age ÷100	-0.00752***
	(0.00218)
Male dummy	-0.0856***
	(0.0107)
White dummy	-0.0576***
	(0.0128)
Rural dummy	0.00477
	(0.0106)
Constant	3.531***
	(0.0403)
Observations	39,784
R-squared	-0.011

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix 6: Comparison of Wealth Index components in 2017 and 2021

Planning group[7]	2017	2021	% change
Household income			
High (4 or 5)	7.956	8.000	0.6%
Neutral (3)	7.980	7.838	-1.8%
Low (1 or 2)	8.167	7.907	-3.2%
Home ownership			
High (4 or 5)	0.628	0.545	-13.2%
Neutral (3)	0.598	0.525	-12.2%
Low (1 or 2)	0.598	0.471	-21.2%
Financial situation			
High (4 or 5)	3.486	3.542	1.6%
Neutral (3)	3.440	3.497	1.7%
Low (1 or 2)	3.441	3.507	1.9%
Financial anxiety			
High (4 or 5)	3.383	3.430	1.4%
Neutral (3)	3.236	3.271	1.1%
Low (1 or 2)	3.153	3.217	2.0%

[7] Planning groups are based on the Likert responses to the variable 'I need to be sure I've considered all the options before I make a purchase decision'. Group sizes for the global sample are as follows: high – 23,402; neutral – 8,710; low – 4,361. Though unevenly distributed, the 95% confidence intervals for each variable are acceptably small.

Appendix 7: Data Definitions of Recurring Variables

Variable	Definition	Comments
Age	Age in years	Respondents aged 15 to 74 answered the survey
Square of age ÷ 100	Age in years, squared, then divided by 100	Used to determine marginally decreasing effects associated with age. Divided by 100 to ensure comparability with other variables
Dummy	Variable taking the value '1' or '0'	Example: 'rural dummy' would take the value 1 if the individual lives in a rural area, and 0 if not
Education level	Highest level of education completed by the individual	Answered on a sliding scale: 1 – 12th grade or less or no diploma 2 – High school or equivalent 3 – Some college, no degree 4 – College degree 5 – Some university, no degree 6 – Bachelor's degree 7 – Master's degree or higher
P1	Planning – "I carefully consider options before a purchase decision"	Answered on a Likert scale: 1 – Strongly disagree 2 – Disagree 3 – Neither agree nor disagree 4 – Agree 5 – Strongly agree
P2	Planning – "I budget carefully when doing the household shopping"	Answered on a Likert scale: 1 – Strongly disagree 2 – Disagree 3 – Neither agree nor disagree 4 – Agree 5 – Strongly agree

[Page de garde]