



Research article

Learning about financial health in Canada

Adam Metzler^{1,*}, Yuhao Zhou¹ and Chuck Grace²

¹ Department of Mathematics, Wilfrid Laurier University, Waterloo, ON, Canada

² Ivey Business School, Western University, London, ON, Canada

* **Correspondence:** Email: ametzler@wlu.ca.

Abstract: This paper applies cluster analysis to eleven (continuous) years' worth of responses to the Canadian Payroll Association (CPA) survey of employed Canadians. The clustering algorithm clearly identifies three distinct groups of respondents. Between-group comparison of response patterns reveals that two of the groups lie on opposite sides of the financial health spectrum, and leads us to label their members “financially stressed” and “financially capable”, respectively. The third group shares characteristics with both the stressed and capable groups, and we label its members as “financially coping”. We find that financial stress is both widespread (one third of all respondents are identified as stressed) and complex (stress is only weakly related to simple demographics such as age or income). From a methodological perspective, an important point is that our use of cluster analysis allows us to generate rigorous insights into financial well-being, without having to measure it directly.

Keywords: financial health; financial well-being; financial stress; cluster analysis; survey data

JEL Codes: C10, G50, I30

1. Introduction

According to the Consumer Financial Protection Bureau (CFPB, 2015), financial well-being is “a state of being wherein a person can fully meet current and ongoing financial obligations, can feel secure in their financial future, and is able to make choices that allow enjoyment of life”. Though this is by no means the universally accepted definition of financial well-being, or the apparently equivalent concepts of financial health or financial wellness, it is eminently similar in spirit to alternative definitions that have appeared in the literature. See Xiao (2016) or Kempson et al. (2017) for careful

delineations between financial well-being and related concepts such as financial capability and financial literacy. It is generally (and implicitly) accepted that financial well-being exists on a spectrum, with financial stress corresponding to the undesirable end of that spectrum.

According to Castro-González et al. (2020), financial well-being is important at both the individual level (as an important determinant of quality of life and general well-being) and the aggregate level (as an important determinant of societal welfare). If it were determined solely by financial resources (e.g. income level and accumulated savings) and/or financial literacy, improving financial well-being would be a “simple” matter of paying people more or educating them appropriately. Unfortunately, a growing body of empirical evidence suggests that while financial resources and financial literacy probably do play a role in determining financial well-being, that role is limited at best. Indeed, the nascent empirical literature (e.g. Vlaev and Elliot, 2014; Prawitz and Cohart, 2016; Collins and Urban, 2020; Strömback et al., 2020; Lee et al., 2020; Casto-González et al., 2020; Iramani and Lufti, 2021) clearly suggests that financial well-being is determined by a complex mix of financial resources, financial behaviours, psychological traits and (possibly) financial literacy. Recent advances notwithstanding, significant gaps still exist in the literature. According to Kempson et al. (2017) “there is a lack of consensus and empirical evidence on how to define financial well-being and how it should be measured. Moreover, there is only fragmented evidence on the key drivers of financial well-being, and how, if at all, the other factors interact to influence it”, and according to Collins and Urban (2020) “financial well-being is generally not well measured”. The present paper helps to fill some of those gaps, both conceptually and methodologically.

This paper applies cluster analysis, a powerful technique in (unsupervised) machine learning, to eleven continuous years’ worth of responses to the Canadian Payroll Association (CPA) survey. The CPA survey covers a variety of topics related to personal financial circumstances, and our use of cluster analysis allows us to generate rigorous insights into financial well-being without requiring us to measure it directly. We find that financial stress (i) has been widespread among employed Canadians for more than a decade and (ii) is not determined by simple demographics such as age, income or geography. Financial stress is not synonymous with low household income, those with high household incomes are not immune from financial stress, and financial stress is not concentrated in a particular age bracket or geographic location. Relative to the existing empirical literature on financial well-being, ours is distinguished by (i) the longevity of our data source and (ii) the statistical methodology that we employ.

1.1. Data

In 2009, the Canadian Payroll Association (CPA) began conducting an annual survey¹ of employed Canadians, asking questions on a variety of subjects related to personal finance. Since that time, more than 35,000 different individuals have completed the survey (some in multiple years). The demographics of the respondents, and the manner in which they responded to questions that were asked in multiple years, are remarkably consistent through time. Because of its longevity and consistency, the resulting data offers a unique opportunity to learn about Canadians’ financial circumstances and how, if at all, those circumstances have changed over time. It is important to note that the CPA survey is a survey of *employed* Canadians, specifically those that are employed by organizations that are

¹ The survey is administered by Framework Partners.

members of the CPA (at the time of this writing more than 10 million Canadians were employed by such organizations).

1.2. Methodology

We perform a separate cluster analysis of each year's survey data. The clustering algorithm arranges respondents into groups (called clusters), in such a way as to ensure that members of the same group are as similar as possible (with respect to their survey responses) and members of distinct groups are as different as possible. In effect, the algorithm is able to identify response patterns that clearly differentiate respondents, regardless of how complex those patterns might be. The optimal number of clusters, as well as the optimal arrangement of respondents into clusters, are determined by the algorithm in a completely objective fashion according to rigorous mathematical criteria. In order to ensure that our results are comparable through time, we classify respondents based on their responses to those questions (which we call the core questions) that were asked in each of the eleven years.

1.3. Key findings

The algorithm consistently identifies three clusters of respondents, differentiated by the manner in which they responded to certain core questions. The typical member of the first group would find it difficult to manage a brief financial setback (specifically, a one-week delay in the receipt of their pay), saves little to none of their income, and places greater emphasis on the monetary aspects of employment (such as salary or benefits) as compared to non-monetary aspects (such as time off or work-life balance). The typical member of the second group would find the setback manageable, saves a much higher proportion of their income and places greater emphasis on the non-monetary aspects of employment. Because they appear to lie at opposite ends of the financial well-being spectrum, we label these groups "financially stressed" and "financially coping", respectively. The typical member of the third group-which we label the financially coping group-lies somewhere between the other two.

One third of all respondents (aggregated over the entire eleven years period) fall into each of the three clusters. The proportion of respondents in each cluster does vary from year to year, but there is no clear trend through time (i.e. it is difficult to reject the hypothesis that the annual variation is simply noise). This is alarming from a policy perspective, as it suggests that financial stress has been widespread among employed Canadians for some time.

We find that financial well-being cannot be explained by simple demographics. In particular it appears to be completely agnostic with respect to geography and gender, and only weakly related to age (younger respondents are marginally more likely to be classified as stressed, but half of the stressed group is over the age of 40 and a quarter is over the age of 50). Low household income does increase the risk of financial stress but does not guarantee it (roughly half of those with household incomes below \$50,000 are classified as stressed, whereas a fifth are classified as comfortable) and those with high household incomes are not immune from financial stress (although half of those with household incomes above \$150,000 are classified as comfortable, a fifth are classified as stressed).

We also investigate responses to non-core questions, i.e. questions that were not asked in every year and responses to which were not used to classify respondents. This exercise served the dual purpose of validating the labels we had chosen for the clusters and developing a deeper understanding of each cluster. This analysis confirms that savings habits and financial resilience are important

components of financial well-being. In addition, it illustrates that (i) financial stress appears to have a direct impact on personal well-being, (ii) budgeting behaviours are not strongly related to financial well-being, (iii) the stressed and comfortable groups can have different motivations for savings, and different ideas about how to improve their financial situation.

1.4. Contributions to existing literature

Excellent guides to the literature on financial well-being can be found in Xiao (2016), Kempson et al. (2017) and Nanda and Banerjee (2021). This paper contributes to this literature, both conceptually and methodologically.

On the conceptual side, our findings confirm and/or provide additional clarity on the roles that various factors play in determining financial well-being. For example, the limited role played by household income in determining cluster membership is eminently consistent with (and therefore supports) the findings of Vlaev and Elliot (2014), Prawitz and Cohart (2016) and Collins and Urban (2020), who find that household income plays a limited role (at best) in determining financial well-being. Indeed, we are unaware of any empirical studies that find household income plays a dominant role in determining financial well-being. As another example, both Prawitz and Cohart (2016) and Iramani and Lutfi (2021) find that financial behaviours are strongly related to financial well-being (among both the US and Indonesian populations), but in both cases the scale used to measure financial behaviours tracks several different behaviours and it is not clear how much of the overall explanatory power can be attributed to each of the individual behaviours. The markedly different savings habits of our stressed and comfortable groups suggest that savings habits are a critical component of financial well-being, and could potentially be responsible for most of the behaviour scales' explanatory power. Further research into the relative importance of different behaviours would be of clear value to financial planners and counsellors seeking to empower their clients to reduce their levels of financial stress. As a final example, the striking difference between our stressed and comfortable groups' abilities to withstand brief financial setbacks suggests that financial resilience is an important component of financial well-being. This appears to be consistent with one of the key findings in Kempson et al. (2017), namely that a latent factor related to financial resilience is one of three latent factors that explain a considerable proportion of observed variation in responses to their financial well-being questionnaire (the authors do not report the proportion of variation explained by each of the three factors) among the Norwegian population. According to Salignac et al. (2019), there are massive gaps in the literature on financial resilience, and our findings suggest that this is a critical avenue for future research.

In addition to clarifying and confirming previous findings, our analysis also generates several new empirical insights into financial well-being. For example, the stressed group places much greater importance on the monetary (as opposed to non-monetary) aspects of employment, relative to the comfortable group. Because it is related to financial well-being, but is unrelated to anything on existing scales, items related to this differential emphasis could potentially be useful additions to future iterations of those scales. As another example, the degree to which financial stress appears to impact overall well-being is not surprising, but does not appear to have been explicitly documented in any of the aforementioned empirical studies.

On the methodological side, an important novelty of our approach is that it allows us to generate rigorous insights into financial well-being, without having to measure it directly (the clustering methodology itself is not novel, but its application in this context is). This contrasts with existing

empirical approaches, which first attempt to measure financial well-being at the individual level (e.g. using a scale), and then use some form of regression analysis to investigate the impact of various factors on those measurements. Notable scales include (but are not necessarily limited to) the Personal Financial Wellness Scale (PFW) developed by Prawitz et al. (2006), the Financial Well-Being (FWB) scale, developed in CFBP (2017), and a scale developed by Kempson et al. (2017) and Kempson and Poppe (2018). We are (most emphatically) *not* suggesting that cluster analysis replace existing scales or scale-based approaches, rather we *are* suggesting that our approach could play a valuable complementary role.

2. Survey and respondents

The CPA survey is a survey of employed Canadians. No efforts are made to contact individuals that are in school, retired or unemployed (that being said, occasionally such individuals do complete the survey). It is important to note, therefore, that the experiences of specific groups could be very different from those of our sample. In the next section we will, where possible, identify groups that might not be covered by our sample and provide references to studies that focus specifically on the financial circumstances of those populations.

With minor variations from year to year the survey consists of ten sections—respondent profile, employment profile, remuneration and financial situation, financial literacy, saving, debt, economic confidence and financial priorities, pay and tax statement perceptions, payroll/tax/deductions/benefits, and survey wrap-up. The length of the survey varies through time, from a minimum of 30 questions in 2010 to a maximum of 57 questions in 2019, with an average of 46 questions per year. On average 3,680 responses were received per year, ranging from a minimum of 2,752 in 2013 to a maximum of 5,629 in 2016. The number of responses in each year is reported in the table below. Outside of the first two years (when it was close to 60%) the completion rate was very close to 90% in every year, ranging from a minimum of 85% to a maximum of 94%.

Table 1. Annual number of responses to the CPA survey.

YEAR	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
#Respondents	2812	3753	2070	3365	2752	3211	3568	5629	4766	5074	4277

Initially there was very little overlap among respondents from distinct years, but that has begun to change in recent years. For instance, the percentage of respondents that were completing the survey for the first time decreased steadily from 97% in 2010 to 74% in 2018. The percentage of respondents that had completed the survey in at least one prior year increased from 3% in 2010 to 26% in 2018, and the percentage that completed in at least two prior years increased from less than 1% in 2011 to 7% in 2018. Very few individuals have completed the survey in three or more years. An important decision is whether or not to include (i) all respondents that completed the survey in a given year or (ii) only those that were completing the survey for the first time. We applied our clustering algorithm

under both assumptions and found that in all but one year there was no difference in our results, in the sense that every individual was assigned to the same cluster under both assumptions².

2.1. Respondent demographics

The demographics of the survey respondents are remarkably stable through time. In what follows we consider the demographics of all respondents, i.e. aggregated over the entire eleven years period. Because the variation across years is so small, the demographics of the aggregate sample of all respondents are eminently representative of the demographics in any given year.

Age. Age demographics are as follows. Notably, each of the three generations (Millennials, Generation X and Baby Boomers) are well represented in the sample. It is important to note that 80% of respondents are between the ages of 30 and 60, so the experiences of specific groups (e.g. post-secondary students, recent post-secondary graduates or senior citizens) could potentially be different from those of our sample. See Gutter and Copur (2011) or Cude et al. (2016) for discussions of the financial circumstances of college students in the US, and DeVaney (2016) for those faced by older individuals.

Table 2. Age demographics of survey respondents, from 2009–2019.

<i><30 Years</i>	<i>30–39</i>	<i>40–49</i>	<i>50–59</i>	<i>>60</i>
14%	26%	27%	26%	7%

Geography. Geographic demographics are as follows. They are comparable to that of the Canadian population as a whole, although residents of Quebec are moderately underrepresented and residents of the prairie provinces are moderately overrepresented.

Table 3. Geographic demographics of survey respondents, from 2009–2019.

<i>British Columbia</i>	<i>Prairies</i>	<i>Ontario</i>	<i>Quebec</i>	<i>Atlantic Canada</i>	<i>Territories</i>
15%	35%	32%	11%	6%	1%

Household Income. The distribution of household incomes across the sample is as follows. The median household income is somewhere between \$80,000 and \$100,000, which is considerably higher than the Canadian median of \$70,000 (according to the 2016 census). To the extent that the Canadian median includes households where all members are either unemployed or retired, this is to be expected. Notably, both lower and higher income (employed) Canadians are reasonably well represented among the sample. For instance, 15% of respondents have household incomes below \$50,000 (well below the Canada median) and 18% have incomes above \$150,000 (well above the Canadian median). That being said, the sample is skewed towards middle- and high-income individuals, so the experiences of low-income

² In 2011, approximately two thirds of all respondents were assigned to the same cluster under both assumptions, and between-cluster variation in responses to important questions was not materially different under either assumption. Results presented in Section 4 correspond to the inclusion of all respondents, not just those completing for the first time.

individuals could potentially be different from those of our sample. See Prawitz et al. (2012) or Nielsen et al. (2016) for discussions of the financial circumstances of low-income individuals in the US.

Table 4. Income demographics of survey respondents, from 2011–2019.

<\$50K	\$50–70K	\$70–100K	\$100–125K	\$125–150K	>\$150K
15%	15%	23%	18%	11%	18%

Gender. Two thirds of respondents are female, one third are male. Males are underrepresented relative to the general population, but not to the extent that we cannot make meaningful inferences about their financial well-being.

Family Status. 51% of the sample is married with children, and 19% are married without children. 11% have children and are not married, and 19% are not married and do not have children. Thus, a majority of the sample (70%) is married, and a majority (72%) has children.

Employment Status. The vast majority of respondents (89%) are employed on a full-time basis. A small minority are employed on a part-time basis (6%), and an even smaller fraction are on contract (2%). The remainder are either self-employed, retired, unemployed or in school. It is important to note that the experiences of our sample could be very different from more vulnerable populations, such as precariously employed or unemployed. See Moscone et al. (2016) for a study of the impact of precarious employment on mental health in Italy.

Occupation. Nearly 40% of the sample is employed in a payroll, accounting or human resources department at their place of work. Beyond this we do not have information on respondents' occupations. We are unsure of how, if at all, the lack of occupational diversity might affect our findings. Because our key findings are broadly consistent with other studies, we suspect that the influence would be negligible. That being said, to the best of our knowledge no studies consider the relationship between occupation and financial well-being, making it an important avenue for future research.

2.2. Core questions

The survey itself varies from year to year, in the sense that the survey from a particular year is not identical to the survey from a different year. For instance if one were to compare the 2010 survey to the 2015 survey, one would find (i) questions that appeared in 2010 but not 2015, (ii) questions that appeared in 2015 but not 2010 and (iii) questions that appeared on both surveys (possibly subject to minor variations in wording or allowable responses). In total we identified 98 distinct questions that were asked at any point during the 11 years.

Our goal is to analyze responses from each year individually, and then make comparisons through time. An important decision is which questions to include when applying the clustering algorithm to a given year's responses. One option is to consider all questions that were asked in a given year, an approach that maximizes the amount of information considered by the algorithm (which is good), but hinders our ability to make direct comparisons between years (which is bad). At the other extreme we could only consider responses to those questions that were asked in every single year, which minimizes information but maximizes comparability. In order to address these issues, we create a list of so-called "core questions" (see the Appendix for details) and, when applying the clustering algorithm to a given

year's responses, we consider the responses to all of the core questions that were asked in that year. This approach strikes an appropriate balance between the two extremes.

Figure 1 illustrates trends in responses to selected core questions through time. For example, the proportion of respondents who report that they would have difficulty managing the one-week delay in their paycheque decreased from 40% to 30% over the eleven-year period, with limited variation between successive years (the fact that this proportion decreased is a positive sign; that it remains near 30% is concerning). With respect to temporal variation, the data exhibited in the figure is representative of responses to core questions. In each case we do see relatively clear long-term trends; beyond that the degree of variation from year to year is limited. The point here is not to uncover or discuss those longer term trends, rather it is to demonstrate a certain temporal stability in the sample, without which comparisons over time would be impossible. In other words, the point of Figure 1 is to illustrate the (considerable) temporal integrity of the data.

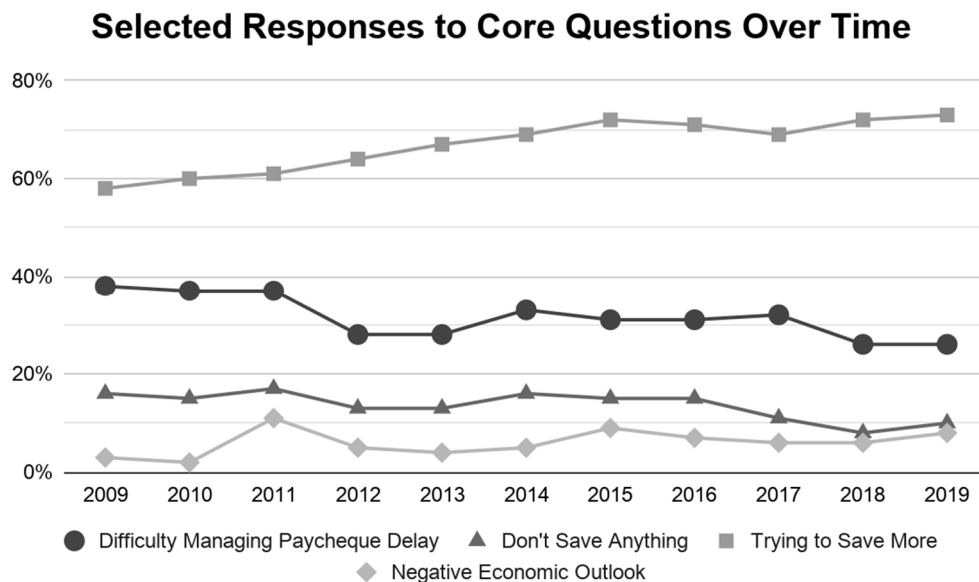


Figure 1. Each curve illustrates the percentage of respondents who selected particular responses to a particular core question (among those who responded to that question). Specifically, circles represent the fraction of respondents who report that it would either be “very difficult” or “difficult” to manage a one-week delay in the receipt of their pay, squares represent the fraction of respondents who report that they are trying to save more than they had in the previous year, triangles represent the fraction of respondents who report that they are not able to save anything in a typical pay period, diamonds represent the fraction who report that they feel the economy in their town or city will either “strongly worse” or “worsen” over the coming year.

One third of the core questions are related to financial well-being, half are related to demographics and the remainder are miscellaneous (e.g. related to financial literacy or economic outlook). Our inclusion of demographics in the list of core questions may surprise some readers, and deserves an explanation. Indeed, when using cluster analysis to investigate a specific concept it might seem more

natural to cluster solely on that concept (i.e. using features that are known to be directly related to that concept) and then investigate demographics of the resulting clusters to check whether or not demographics seem to play a role. Our reason for including the demographic variables in the clustering algorithm is simple—our original mandate was to conduct a rigorous exploratory analysis of the CPA data (as opposed to conducting a study of financial well-being in Canada) and for this reason we wanted to feed as much information as possible into the clustering algorithm. It was by no means predetermined that we would end up classifying respondents based on their position on the spectrum of financial well-being. In this sense, their location on that spectrum emerged organically as the most informative way to distinguish respondents, as compared to simple demographics such as age or household income. Thus, and in hindsight, our approach implicitly allowed us to place financial health and demographics on an even footing, and determine which was the more powerful differentiator. We feel that this is a powerful insight and wanted to make sure it came through in this paper. As a robustness check we did repeat all of our analysis when demographics were excluded, and none of our key findings were affected.

3. Clustering algorithm

Cluster analysis is a powerful tool for exploratory analysis of high-dimensional data, having the potential to identify complex features or patterns in large data sets that would otherwise be invisible to the naked eye. See (Hinton et al., 1999) for an excellent introductory treatment. Loosely speaking, the goal of any cluster analysis is to arrange a large number of objects into groups in such a way as to ensure that (i) members of the same group are as similar as possible and (ii) members of different groups are as different as possible, with respect to some numerical measure of similarity. In the present context we measure the similarity between two individuals via the coincidence between their overall survey responses; for instance if two individuals provided similar, but not necessarily identical, responses to 80% of the survey questions, the similarity score between those individuals would be high (conversely, their dissimilarity score would be high). An important point is that the arrangement of individuals into clusters is carried out in a completely objective fashion by the algorithm, once the user has specified the measure of similarity. The user does not tell the algorithm to place more weight on responses to particular questions, or pay more attention to particular demographic variables. In this sense, the arrangement of individuals into groups (clusters) is completely objective and data-driven. It is also possible to let the data speak on the issue of the ideal number of clusters, i.e. the number of clusters that minimizes within-group variation while maximizing between-group variation.

We run a separate cluster analysis on each year's response data. In each year, we identify those questions from the core list that appeared in that year, and conduct the analysis on the responses to those questions. In order to complete the analysis we must choose (i) which respondents to include in a given year, (ii) how to define the distance between two respondents, based on their responses to the core questions, (iii) how to arrange a given year's respondents into a fixed number of clusters and (iv) how to determine the appropriate number of clusters. We use the "cluster" package in R for all of our analysis.

3.1. Distance between respondents

A critical decision is how to define the distance between two respondents in a given year. Mathematically, an individual's responses to the core questions are summarized by a 15-dimensional vector containing both categorical and ordinal variables. In some cases, an individual failed to answer one or more questions, which means that some components of an individual's vector could be missing. In addition, most questions allow for "Don't Know", "Not Applicable" or "Prefer Not to Respond" options; because such responses are uninformative, we treat them as missing values.

Having created a vector of responses for each individual in a given year, we must then decide how to measure the difference between two individuals' responses. If all the components of the vectors contained numerical values this would be straightforward using, say, Euclidean distance. The components of our vectors contain both categorical and ordinal data, and could potentially be missing, which means that Euclidean distance is not appropriate. To this end we use Gower distance (Gower, 1971), which is a standard method for computing distances between vectors whose components are of mixed data types and subject to missing values. See Madhulatha (2011) or Xu and Tian (2015) for more details.

In the present context, the Gower distance between two respondents is the average dissimilarity of their responses, with the average is taken over all questions for which both respondents provided answers (i.e. if at least one respondent did not answer, or selected "don't know" or "prefer not to answer", the question is not included in the calculation). The dissimilarity between responses to a given question depends on whether or not the response variable is categorical (in which case the dissimilarity is zero if the respondents selected the same category and one otherwise) or ordinal (in which case the dissimilarity is the relative difference between the responses). More formally, and borrowing notation from Foss et al. (2018), consider distinct respondents x and y , and some question j . If both respondents provided informative responses to question j (i.e. both answered the question and neither selected "prefer not to respond") then we denote the dissimilarity between their responses to question j as $d_j(x, y)$ and define it as follows:

If question j is categorical then $d_j(x, y) = 1$, if both x and y selected the same response, and $d_j(x, y) = 0$ otherwise. For example, two respondents reporting different family statuses would be assigned a dissimilarity of 1 (regardless of what their particular statuses are) and two respondents reporting the same status would be assigned a dissimilarity of 0.

If question j is ordinal then $d_j(x, y) = |x_j - y_j|/r_j$, where x_j and y_j are the responses selected by x and y , respectively, and r_j is the maximum possible difference between two responses to question j . For example, if respondents x and y selected "disagree" and "neither", respectively, for the paycheque delay question then the dissimilarity between their responses would be $|2 - 4|/6 = 1/3$. If instead they had selected "strongly agree" and "somewhat disagree" the dissimilarity would be $|7 - 3|/6 = 1/2$, which is larger. Note that one must first convert the ordinal responses to some numerical scale (which is a straightforward exercise).

The final step is to combine the dissimilarities between individual questions into an aggregate dissimilarity between respondents x and y , which we denote by $d(x, y)$. To this end we compute the average value of $d_j(x, y)$, over all questions j for which it is well-defined (i.e. for which both respondents answered, and neither selected "prefer not to respond"). For example, if both respondents provided informative answers to all 12 questions then we would have $d(x, y) = (d_1(x, y) + \dots + d_{12}(x, y))/12$. If instead, and as a pair, they only provided informative responses to

questions 1, 4, 7 and 10, the aggregate dissimilarity between the two respondents would be $d(x, y) = (d_1(x, y) + d_4(x, y) + d_7(x, y) + d_{10}(x, y))/4$.

In order to compute the dissimilarity between each pair of respondents, we use the *daisy* function in R, with the *metric* argument set to *gower*. No other arguments are passed to *daisy*.

3.2. Clustering algorithm

We use the Partitioning Around medoids (PAM) algorithm to arrange a particular year's respondents into a given number of clusters (choosing the optimal number of clusters is discussed in the next section). The basic idea is to identify a "most representative" member of each cluster, called the cluster's *medoid*. The medoid is chosen such that the distance between it, and all other members of its cluster, is as small as possible. The PAM algorithm identifies medoids iteratively as follows; see (Kaufman and Rousseeuw, 1987) or (Reynolds et al., 2006) for further details.

Suppose for the moment our goal is to arrange respondents into four clusters. The PAM algorithm would then proceed as follows.

1. Randomly select four respondents and deem each one the medoid of a different cluster. For example, we might randomly select respondents 2071, 13, 875 and 1042 to be the initial medoids. We write $m_1 = 2071$ to indicate that respondent 2071 is the medoid of the first cluster, etc. for $m_2 = 13$, $m_3 = 875$ and $m_4 = 1042$.

2. For each of the non-medoid respondents, compute the distance between that respondent and each of the four medoids, and assign that respondent to the cluster associated with the closest medoid. For example if $d(1, m_1) = 0.94$, $d(1, m_2) = 0.25$, $d(1, m_3) = 0.63$ and $d(1, m_4) = 0.17$, then the first respondent would be assigned to cluster 4. In general, respondent x is assigned to that cluster k for which $d(x, m_k)$ attains its minimum value. We write S_k to denote the set of all respondents that are assigned to cluster k .

3. For each member of a given cluster, compute the average distance between (i) that member and (ii) every other member of that cluster. For example, suppose that respondent 473 was assigned to cluster 1 in the previous step. Here, we would compute the average value of $d(473, x)$, as x ranges over all respondents that were assigned to cluster 1 (i.e. over all x in S_1); call the result $a(473)$. The result is a measure of how close that member is to the "true centre" of the given cluster. Intuitively, if $a(473)$ is small then respondent 473 is relatively similar to most members of its cluster, and is therefore close to the "true centre" of the cluster; if it is large then respondent 473 is distant from a substantial number of cluster members, and is therefore not close to the "true centre". Now replace the current medoid of the given cluster with that cluster member having the lowest average distance from the other members. For example, the new medoid of cluster 1 would be that value of x in S_1 for which $a(x)$ attains its minimum value. If, for example, $a(x)$ is minimized for $x = 78$, then respondent 78 would become the new medoid for cluster 1 and we would set $m_1 = 78$. Repeat this exercise for every cluster.

4. Iterate Steps 2 and 3 until the allocation of respondents to clusters does not change.

In order to implement the PAM algorithm in R for a fixed number of clusters, we pass the dissimilarity matrix discussed in Section 3.1 into the *pam* function in R. Despite its iterative nature and the sheer number of calculations that must be carried out, implementing the PAM algorithm on a single year's data only requires a few seconds of CPU time.

3.3. Which respondents to include

Although the overall completion rate is consistently in excess of 90%, many respondents that completed the survey selected responses such as “prefer not to answer” or “don’t know” to many of the core questions. For example, in an average year only 40% of all respondents provide informative answers to all of the core questions, whereas 82% of respondents provided informative responses to at least 70% of the core questions.

If a respondent provides an excessive number of uninformative responses, it can be difficult to accurately assess their distance from other respondents (recall that the Gower distance simply ignores such responses). As such, we decided to exclude respondents that provided uninformative responses to more than half of the core questions, i.e. we only included those that provided informative answers to at least 50% of the core questions. In an average year, this meant that we included 90% of all respondents.

As a robustness check we experimented with the minimum threshold³ and found that (i) broadly speaking, key insights were unaffected by the threshold and (ii) the 50% threshold gave us the clearest distinctions between the three clusters. As a further robustness check we also investigated (i) whether or not members of a particular cluster tended to provide a higher proportion of informative responses and (ii) whether or not members of a particular cluster were more likely to provide uninformative responses to a particular question, and did not find any clear patterns in either case.

3.4. Number of clusters

Recall that the ultimate goal is to arrange respondents into groups in such a way as to ensure that (i) members of the same group are as similar as possible and (ii) members of distinct groups are as different as possible. In other words, the goal is to simultaneously minimize within-group variation and maximize between-group variation. Generally speaking, increasing the number of clusters will reduce within-group variation but could either increase or decrease between-group variation (decrease being more likely if the number of existing clusters is large, increase being more likely if the number of existing clusters is small), so adding an additional cluster is only warranted if it does not blur the distinction between groups. A variety of criteria for identifying the ideal number of clusters have been proposed in the literature (Bezdek and Pal, 1998). Of those, the so-called silhouette width has been shown to perform well in a variety of contexts (Arbelaitz et al., 2013) and we elect to apply it here. See (Rousseeuw, 1987) for a description of the method, which we now describe at an intuitive level.

Assume for the moment that the number of clusters, and arrangement of respondents into clusters, has been determined. The degree to which a given respondent “fits in” with members of a given cluster (not necessarily the cluster to which the given respondent belongs) can be measured by computing the average distance between the given respondent and all members of the given cluster. The lower this average distance, the better the respondent fits in. Ideally every respondent fits in best with the cluster to which it belongs, but this need not be the case. The respondent’s “neighbouring cluster” is defined as that cluster, among those to which it does not belong, with which it best fits. The silhouette value of a given respondent is defined as the relative difference between (i) the average distance between the respondent and all other members of its cluster and (ii) the average distance between the respondent

³ Specifically, we ran the same analysis but only included respondents that provided informative responses to at least X% of the core questions, for various values of X between 50 and 100.

and all members of its neighbouring cluster. Intuitively, if a respondent's silhouette value is large, then the respondent has much more in common with members of its own cluster than it does with members of any other cluster, which is the ultimate goal of the analysis.

Given the number of clusters, and arrangement of respondents into clusters, the silhouette width is defined as the average silhouette value of all respondents. Large silhouette widths suggest effective arrangements, in the sense that within-group variation is low while between-group variation is high. We use silhouette width to determine the ideal number of clusters as follows. For each value of K between 3 and 7, we use the PAM algorithm to arrange respondents into K clusters and compute the silhouette width of the given arrangement. The ideal number of clusters is then identified as that value of K which produces the largest silhouette width. Table 6 below reports the results of this exercise. In ten of the eleven years, this approach identifies three as the ideal number of clusters, and in the remaining year three clusters yields a silhouette value very close to the maximum (obtained with four clusters). Because the results are so consistent through time, we conclude that three clusters are ideal.

Table 5. This table reports the silhouette width obtained by applying the PAM algorithm with the indicated number of clusters, in the indicated year. For example, applying the algorithm to the 2012 data with 5 clusters leads to a silhouette width of 0.053.

#of clusters	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Average
3	0.117	0.128	0.062	0.064	0.058	0.104	0.094	0.073	0.075	0.103	0.096	0.088
4	0.104	0.120	0.052	0.054	0.052	0.089	0.097	0.069	0.056	0.075	0.081	0.077
5	0.087	0.110	0.051	0.053	0.052	0.087	0.080	0.058	0.051	0.060	0.062	0.068
6	0.064	0.090	0.046	0.050	0.050	0.082	0.077	0.063	0.048	0.070	0.063	0.064
7	0.053	0.076	0.043	0.049	0.044	0.068	0.083	0.057	0.043	0.070	0.055	0.058

Table 6. This table illustrates certain features of each cluster's medoid, for the 2018 data. In some cases, the medoid did not answer all five of the questions appearing in the table, in which case we selected the closest cluster (in the sense of Gower distance) member that did answer all five questions.

Question	Stressed (Male)	Coping (Female)	Comfortable (Male)
Paycheque Delay	Difficult	Somewhat Difficult	Manageable
How Much Saved	1–5%	1–5%	5–10%
Household Income	\$100K–\$110K	\$250K–\$500K	\$110K–\$120K
Age	30–39	40–49	30–39
Most Important	Higher Wages	Work/Life Balance	Work/Life Balance

We conclude with a brief explanation of why we did not consider two clusters. When using PAM with two clusters we found that the algorithm placed an inordinate amount of weight on the “save more” question. For instance, when requesting two clusters from the 2010 data the algorithm placed all respondents that had selected “yes” into one group; the second group consisted of all those who

selected “no”. Delineating respondents based on their responses to a single question is not insightful, which led us to avoid using two clusters.

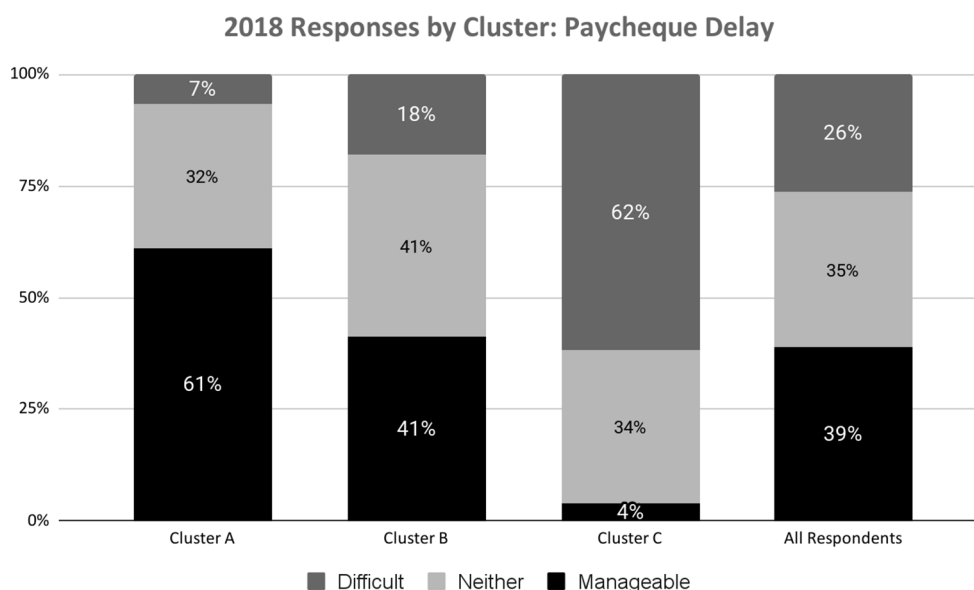


Figure 2. This figure illustrates between-cluster variation in responses to the paycheque delay question. For example, 61% of Cluster A would find the delay manageable (in the sense that they selected “very manageable” or “manageable”), 7% would find it difficult (in the sense that they selected “very difficult” or “difficult”) and 32% were non-committal (in the sense that they selected “somewhat difficult”, “neither manageable nor difficult” or “somewhat manageable”). Note that Clusters A, B and C will eventually be identified as financially comfortable, financially coping and financially stressed, respectively.

3.5. Sensitivity

For each of the eleven data sets, we use the PAM algorithm with Gower distance to arrange respondents from a given year into three clusters. The algorithm begins with an initial assignment of the medoids, which is simply a random guess. It is important to determine how sensitive the resulting arrangement is to the initial guess, and to this end we repeated the arrangement several times with different (and randomly selected) initial selections for the medoids. In every case, the medoids that were ultimately identified by the algorithm were identical; in other words, the algorithm appears to converge to the exact same solution regardless of the initial guess (i.e. it is robust with respect to the initial guess).

It is also important to determine whether or not the number of clusters identified is an artefact of the algorithm used to produce the clusters. To this end we also implemented the k-modes algorithm which, strictly speaking, treats all variables as categorical (and is therefore implicitly less desirable than PAM with Gower distance). The silhouette criteria continued to identify three as the optimal number of clusters, and the interpretation of those clusters was identical. This finding suggests that (i) our identification of three clusters and (ii) the interpretations of those clusters, is robust against

alternative approaches. As a further exercise we implemented the k-means algorithm and obtained nonsensical results. The silhouette criteria suggested either two or three clusters, but there were no clear features distinguishing those clusters. Since k-means treats all variables as continuous (and therefore does not seem appropriate for our data), it is not surprising that it produced nonsensical results here.

4. Analysis

4.1. Interpreting the clusters

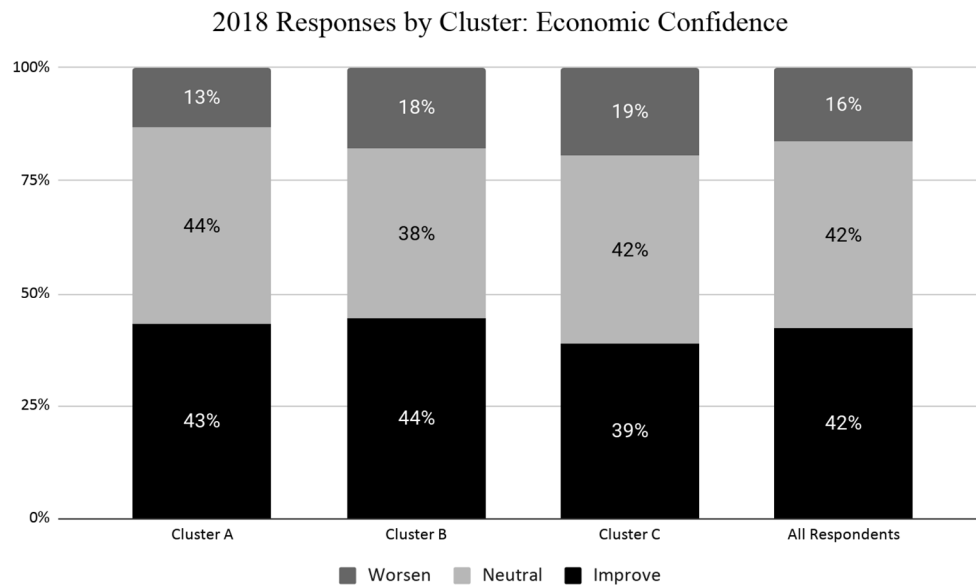


Figure 3. This figure illustrates between-cluster variation in responses to the economic confidence question. For example, 43% of Cluster A expect their local economy to improve (in the sense that they selected “strongly improve” or “improve”), 13% expect it to worsen (in the sense that they selected “strongly worsen” or “worsen”) and 44% were non-committal (in the sense that they selected “somewhat improve”, “neither improve nor worsen” or “somewhat worsen”). Note that Clusters A, B and C will eventually be identified as financially comfortable, financially coping and financially stressed, respectively.

The clustering algorithm described in Section 3 divides respondents from a particular year into three (non-overlapping) groups, and assigns each group an arbitrary label (e.g. A, B or C). The algorithm does not give the user any direct insights into common characteristics shared by members of a particular group, nor does it provide any insights into characteristics that distinguish members of two different groups. In other words, the clustering algorithm itself does not directly provide answers to questions such as “what do respondents in Cluster A have in common?” and “what distinguishes respondents in Cluster A from respondents in Cluster C?” We will eventually attach a meaningful label (stressed, coping or comfortable) to each cluster, but in order to clearly document the process that we

used to construct these labels we leave them arbitrary for now (recall that it was not predetermined that clusters would be distinguished by their location on the financial well-being spectrum).

In order to understand the defining features of each cluster, we consider the between-cluster variation in responses to the core questions. For example, Figures 2 and 3 demonstrate between-cluster variation in 2018 responses to the paycheque delay and economic outlook questions, respectively. The considerable between-cluster variation observed in Figure 2 suggests that financial resilience (as measured by an ability to deal with a one-week delay in the receipt of one's pay) is a defining feature of the clusters, whereas the insignificant variation observed in Figure 3 suggests that economic optimism is not.

In each year, 3 of the non-demographic core questions exhibit variation comparable to that in Figure 2. We use responses to these questions to identify the distinguishing characteristics of each cluster. The other 5 non-demographic core questions either exhibit limited variation comparable to that observed in Figure 3, or exhibit variation that is inconsistent across years. In either case, responses to these questions do not help us learn about the distinguishing features of each cluster.

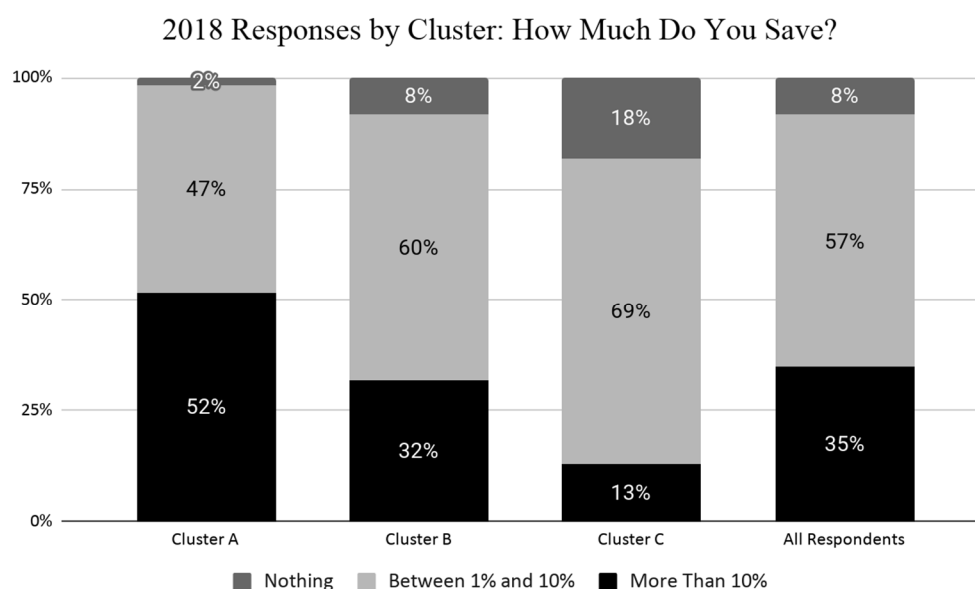


Figure 4. This figure illustrates between-cluster variation in responses to the how much do you save question. For example, 52% of Cluster A saves more than 10% of their paycheque, on average, 47% saves somewhere between 1% and 10%, and 2% don't save anything. Note that Clusters A, B and C will eventually be identified as financially comfortable, financially coping and financially stressed, respectively.

In each year, the greatest between-cluster variation is observed in the same set of 3 questions-paycheque delay, how much do you save, and what is most important. Between-cluster variation in the 2018 responses to these questions are illustrated in Figures 2, 4 and 5, respectively. Inspecting the figures, we see that members of cluster A tend to be financially resilient and, perhaps unsurprisingly, tend to save a relatively high fraction of their income. Indeed only 7% of those in cluster A would find it difficult to manage the paycheque delay and 98% save at least some of their income. By marked contrast, members of cluster C tend to be financially vulnerable and have the

greatest difficulty saving. Indeed only 4% of cluster C would find the brief delay manageable, and 18% are not able to save any of their income. Figure 5 illustrates that members of cluster A tend to place greater emphasis on non-monetary aspects of employment (work/life balance in particular, although that is not illustrated in the figure), which suggests that they are largely comfortable with their financial situation. By contrast, members of cluster C are much more likely to place greater emphasis on monetary aspects (higher wages in particular, although that is not illustrated in the figure), suggesting that they are less comfortable with their financial situation.

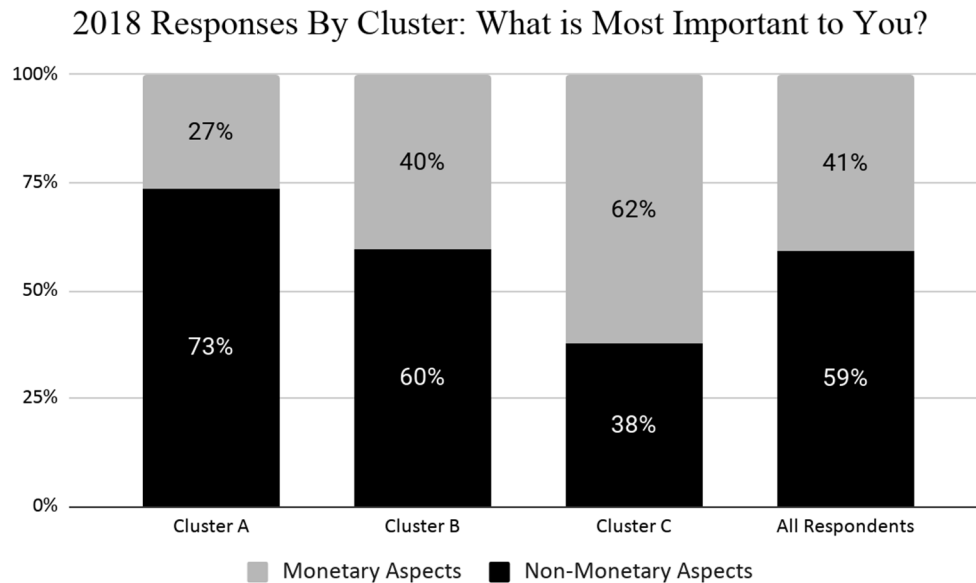


Figure 5. This figure illustrates between-cluster variation in responses to the what is most important question. For example, 73% of Cluster A selected one of the non-monetary options (e.g. better work-life balance, more time off) and 27% selected one of the monetary options (e.g. higher wages or better retirement funding). Note that Clusters A, B and C will eventually be identified as financially comfortable, financially coping and financially stressed, respectively.

It seems clear that members of clusters A and C tend to lie at opposite ends of the financial wellness spectrum, whereas members of cluster B tend to lie somewhere in between. In other words, clusters seem to be distinguished by their location on the financial well-being spectrum. The pattern observed in 2018 is observed in every other year. In particular, in every single year we find that:

1. The cluster that is most resilient (as measured by the ability to deal with the delay in pay) tends to have the strongest savings habits, and places greater emphasis on the non-monetary aspects of employment. We will henceforth refer to its members as financially comfortable.
2. The cluster that is least resilient has the weakest savings habits, and places greater emphasis on the monetary aspects of employment. We will henceforth refer to its members as financially stressed.
3. The third cluster lands squarely between the other two in terms of resilience, savings habits, and the relative emphasis on monetary versus non-monetary aspects of employment. We will henceforth refer to its members as financially coping.

Although the distinguishing features of each cluster remain remarkably consistent through time, their precise characteristics are subject to some variation. For example, Figure 6 illustrates that the stressed cluster is always the least resilient (as measured by self-reported ability to deal with the paycheque delay), but in some years it appears more resilient than in others. Similar phenomena are observed with respect to savings, and the relative emphasis on monetary versus non-monetary aspects of employment.

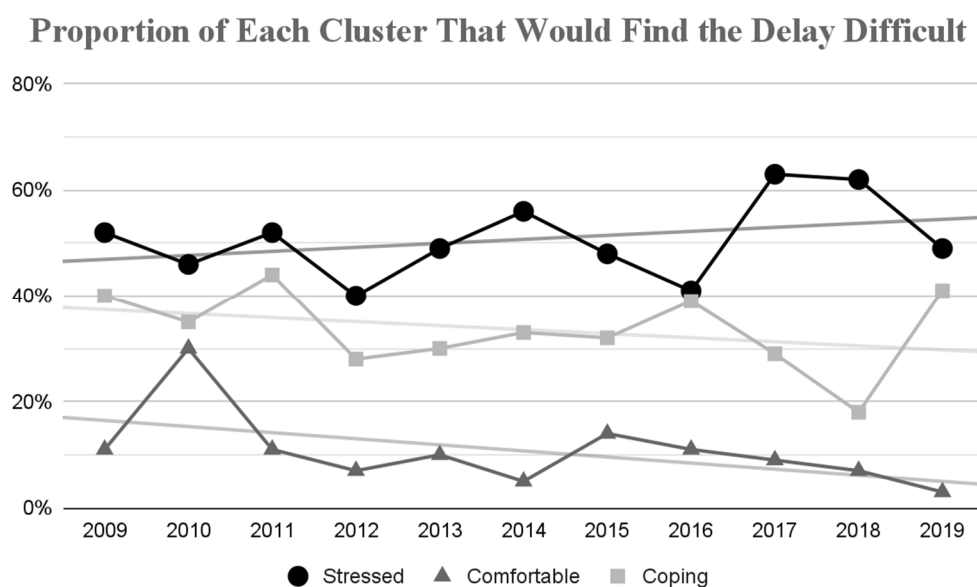


Figure 6. This figure illustrates the proportion of each cluster that (i) answered the paycheque delay question and (ii) selected either “very difficult” or “difficult”. For example, in 2011, among all respondents in the stressed cluster that responded to the paycheque delay question, 52% selected one of these two options. By comparison, the proportions were 44% for the coping cluster and 11% for the comfortable cluster.

The relative size of each cluster also fluctuates through time, to a certain degree. For example, the fraction of all respondents that were considered by the clustering algorithm and classified as financially stressed ranged between a minimum of 23% (in 2017) and a maximum of 51% (in 2009), with an average value of 35%. The size of the stressed cluster also appears to be decreasing at a modest rate over time; if S_t denotes the proportion of respondents that are classified as stressed in year t , then the average value of $S_{t+1} - S_t$ is -2.49% . The size of the comfortable cluster ranges between 24% (2017) and 46% (2019), with an average level of 34% and an average annual change of $+1.72\%$. For the coping cluster the extreme sizes are 20% (2009) and 54% (2017), the average size is 33% and the average annual change is $+0.77\%$. The degree to which these fluctuations, and the apparent trends, are due to changes in actual financial circumstances or fluctuations in precise cluster characteristics as described in the previous paragraph, is impossible to determine.

Before moving on, we briefly discuss responses to the non-demographic core questions that were (considerably) less informative than those discussed above. First, there is essentially no between-cluster variation in responses to the economic outlook and retirement amount questions. This is somewhat

surprising, as Strömback et al. (2020) find that optimism explains a considerable degree of observed variation in financial well-being scores among Swedish university students. Second, and with respect to the paycheque accuracy question, there is no discernible difference between the coping and comfortable groups; the stressed group is slightly less likely to be confident in the accuracy of their pay and slightly more likely to be noncommittal. Third, the only clear difference in responses to the debt free question is observed in the fraction of respondents selecting “never”. In all but one year (2016) the stressed group is most likely to select this option (averaging 16% across years) and in all but one year (2015), the comfortable group was least likely to select this option (averaging 6% across years). Fourth, in an average year 81% of the stressed group reports that they are trying to save more than they did in the previous year; with the exception of 2013 (when it was abnormally low at 24%) this fraction never fell below 78%. This clearly suggests that the financially stressed group is self-aware, in the sense that they recognize that saving more would improve their financial situation. For both of the other groups, the fraction reporting that they were trying to save more fluctuates wildly from year to year.

4.2. Demographics

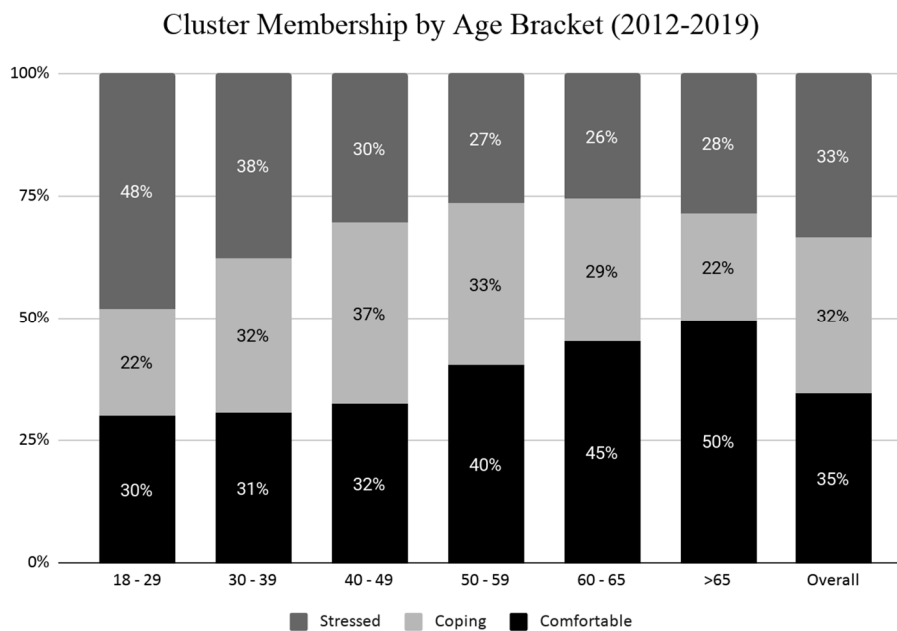


Figure 7. This figure illustrates the distribution of each age bracket across clusters, since 2012 (when the survey began using different age brackets). For example, of all 18–29 years old that completed the survey during this period, 30% were classified as comfortable (in the year they completed the survey), 22% were classified as coping and 48% were classified as stressed.

Figure 7 illustrates the relationship between cluster membership and age. For example, 48% of respondents that were 18–29 when they completed the survey, were classified as stressed in the year that they completed the survey. By contrast, only 26% of those aged 60–65 was classified as stressed. We clearly see that older individuals are more likely to be comfortable, and younger individuals are

more likely to be stressed. That being said, financial stress is most emphatically not concentrated in a single age bracket—at least one quarter of every, but no more than half of any, age bracket is financially stressed. Figure 7 clearly suggests that age is related to, but does not completely determine, financial well-being. This finding is broadly consistent with Collins and Urban (2020), who find that, in and of itself, age is a statistically significant, but not overly important, predictor of FWB scores among the general US population. When it is the only explanatory variable in their regression model age does appear to be statistically significant at high confidence levels, but its explanatory power is limited in the sense that it only explains 6% of observed variation in scores; upon controlling for other factors, however, age begins to lose its statistical significance⁴. Prawitz and Cohart (2016) find that age is not a significantly significant predictor of PFW scores among the general US population, and that after controlling for a wide variety of other factors (including financial behaviour scores, income, accumulated savings and debt-to-income ratio), age (or gender) cannot explain further variation of PFW scores. A possible explanation for the discrepancy between our finding and theirs, is that age is correlated with other factors in their model (e.g. financial behaviour scores).

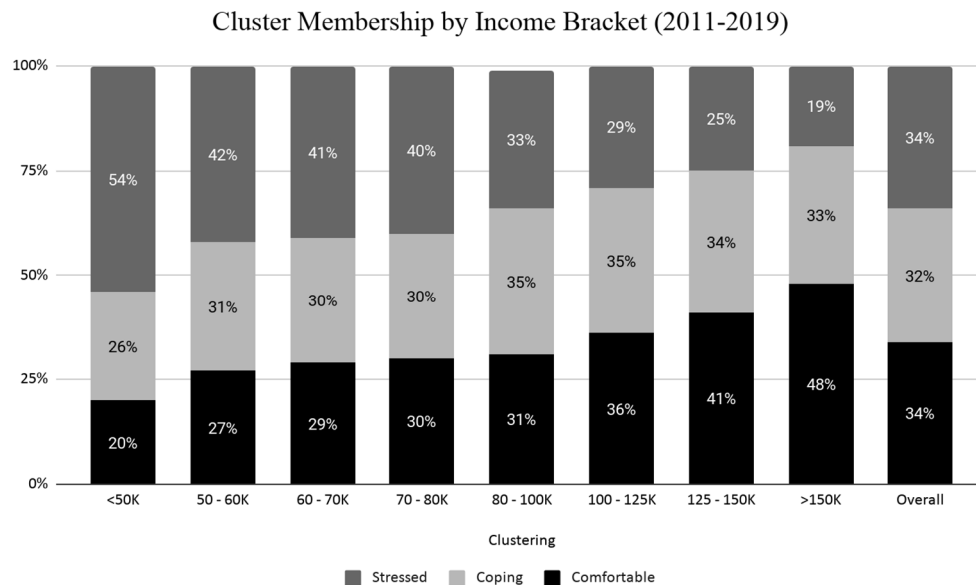


Figure 8. This figure illustrates the distribution of each income bracket across clusters, since 2011 (when the survey began using different income brackets). For example, of all respondents who completed the survey during this period and reported a household income of less than \$50,000, 20% were classified as comfortable (in the year they completed the survey), 26% were classified as coping and 54% were classified as stressed.

Figure 8 illustrates the relationship between cluster membership and household income. The lower one's income is, the more likely it is that they are financially stressed. But by no means is financial stress synonymous with low income (20% of those earning less than \$50,000 are classified

⁴ Even after controlling for an extremely wide range of factors, it appears as though being over 62 leads to significantly better wellness scores, but this might simply be due to the fact that the FWB scale uses different weights for those aged 62 and older.

as comfortable), and those with high incomes are not immune from financial stress (19% of those earning more than \$150,000 are classified as financially stressed). Household income here plays a definite, but limited, role in determining financial well-being. This finding is eminently consistent with existing literature, which is essentially unanimous on the definite, but limited, role that financial resources play in determining financial well-being. For example, in their survey of young workers and families with young children in the UK, Vlaev and Elliot (2014) find a weak correlation (magnitude less than 20%) between annual incomes and (seven-point Likert) responses to the question “how satisfied are you with your overall financial circumstances?”, while Prawitz et al. (2006) find that after controlling for Financial Management Behavior Scale (FMBS) scores and locus of control, annual income level and accumulated savings are unable to explain further variation in PFW scores among the general US population. Others have found evidence of a stronger, but still limited, influence of financial resources. For example, the CFPB (2017) finds a moderate correlation (magnitude nearly 40%) between annual household income and FWB scores, and after controlling for financial literacy, subjective well-being and a wide variety of demographic variables, Collins and Urban (2020) find that financial resources (including, but not limited to, income level and accumulated savings) explain an additional 20% of variation in FWB scores. Finally, Iramani and Lutfi (2021) find a moderate correlation (magnitude approximately 40%) between financial well-being (according to the scale developed in that paper) and financial status (as reflected by monthly income and net wealth) among the Indonesian population.

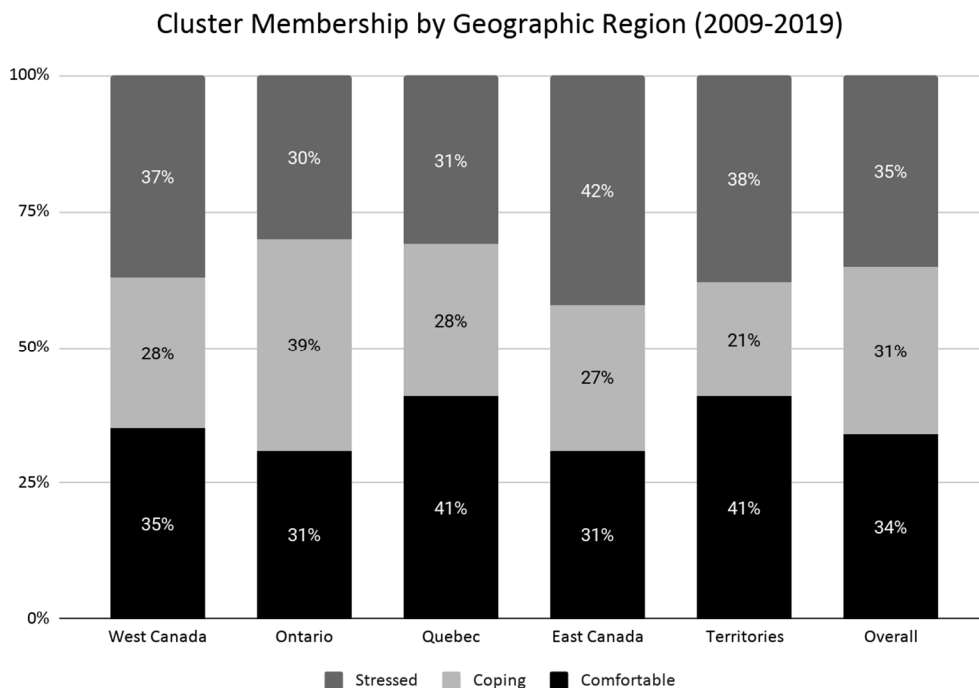


Figure 9. This figure illustrates the distribution of different geographic regions across clusters. For example, of all respondents who completed the survey in any year and resided in Western Canada (Manitoba, Saskatchewan, Alberta, British Columbia) 35% were classified as comfortable (in the year they completed the survey), 28% were classified as coping and 37% were classified as stressed.

Other than age or income, no demographic variable has a notable influence on cluster membership. That being said, males are somewhat more likely to be either stressed or comfortable⁵, Eastern Canadians are marginally more likely to be financially stressed and residents of Quebec are marginally more likely to be financially comfortable (Figure 9), and single individuals are somewhat more likely to be financially stressed than married individuals (Figure 10). Cluster membership is essentially identical across employment types (in particular respondents that are employed on a full-time basis are no more or less likely to be stressed or comfortable, as compared to respondents that are employed on a part-time basis) and organization size. Our findings with respect to gender are broadly consistent with existing studies, for example both Collins and Urban (2020) and Prawitz and Cohart (2016) find that females tend to have higher financial well-being scores, but that gender is not an overly important predictor of those scores. Finally, our findings with respect to marital status are consistent with those of Collins and Urban (2020), who find that married individuals tend to have higher FWB scores, but that marriage is not an overly important predictor of those scores.

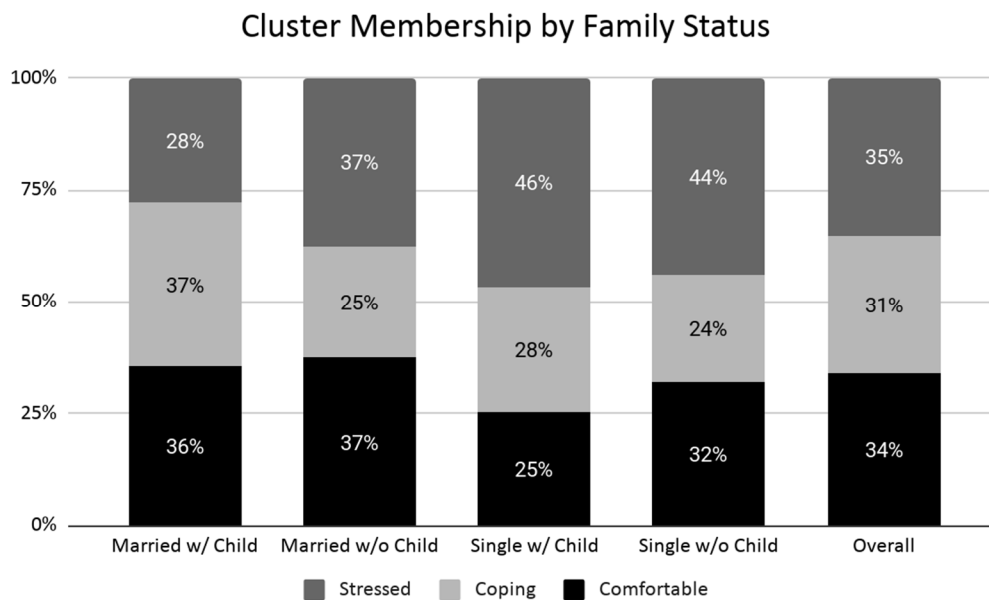


Figure 10. This figure illustrates the distribution of family status across clusters. For example, of all respondents who completed the survey in any year and were married with children (at the time they completed the survey), 36% were classified as comfortable (in the year they completed the survey), 37% were classified as coping and 28% were classified as stressed.

4.3. Non-core questions

In any given year, the core questions are a relatively small fraction of the total survey. Having assigned respondents to clusters and developed a basic understanding of each cluster, our next task is

⁵ 39% of males were categorized as stressed in the year that they completed the survey, 41% were classified as comfortable. The corresponding proportions for females were 33% and 32%, respectively.

to investigate how the clusters responded to non-core questions. This serves the dual purpose of (i) validating the labels that we have given the clusters and (ii) generating additional insights about each cluster. It is important to keep in mind that the clustering algorithm was not provided with responses to non-core questions; when deciding which group a particular respondent belonged to, the algorithm was completely unaware of how that respondent answered the non-core questions.

Analysis of the non-core questions confirms that both savings habits and financial resilience are important components of financial well-being. In particular the stressed group would have the greatest difficulty raising \$2,000 on short notice (Figure 11), are most likely to spend all or more of their net pay (Figure 12) and are most likely to report that their financial position deteriorated in the previous year (Figure 13). In addition, the stressed group tends to be furthest from its retirement goals, most likely to report an increasing debt load, most likely to have credit card debt and least likely to successfully achieve savings goals.

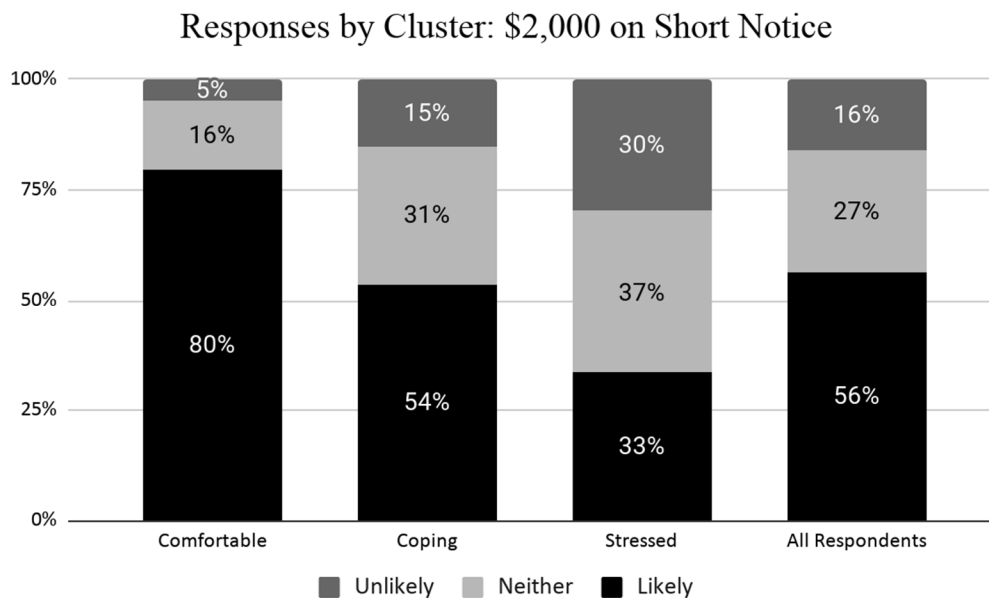


Figure 11. This figure illustrates between-cluster variation in responses to the question “how likely are you to come up with \$2,000 if an emergency arose within the next month”. Responses are on a 7-point Likert scale from “very likely” to “very unlikely”. For example, 80% of the comfortable group selected “very likely” or “likely”, 16% selected “somewhat likely”, “neither”, “somewhat unlikely” and 5% selected “unlikely” or “very unlikely”.

Our analysis here reinforces the critical role that savings habits play in determining financial well-being, and helps to refine some of the findings in other studies. For example, after controlling for age, financial literacy and a variety of demographic factors, Collins and Urban (2020) find that income and accumulated savings explain an additional 20% of observed variation in FWB scores among the general US population, but do not separately consider the relative importance of the two; our findings suggest that savings are at least as important as income. It is also interesting to note that after controlling for FMBS scores (of which savings habits are a component) and locus of control, Prawitz and Cohart (2016) find that income and accumulated savings explain very little additional

variation in observed PFW scores; our findings suggest that savings *habits* could play a very different (and potentially much more important) role than *accumulated* savings, in determining financial well-being.

Of particular note is the \$2,000 question, which formed the basis for an excellent study by Lusardi et al. (2011) of financial fragility in the general US population. Interestingly, the authors of that study found that “financial fragility ... is not just a poor person’s problem: a material fraction of the solidly middle class is pessimistic about their ability to come up with \$2,000 in a month”, which is entirely consistent with our findings on (i) the relationship between financial well-being and financial resilience and (ii) the limited role that income plays in determining financial well-being. Importantly, Lusardi et al. (2011) find that recent financial shocks (e.g. loss of job or wealth, which are not visible in the CPA data) increase, but do not fully explain, the likelihood that an individual is financially fragile. Salignac et al. (2019) contend that financial resilience is an underdeveloped topic in the academic literature and suggest that enhanced financial resilience could lead to improved financial well-being. Our findings in this paper strongly support their suggestion which, in light of their contention, indicates that further research into the determinants of financial resilience is needed.

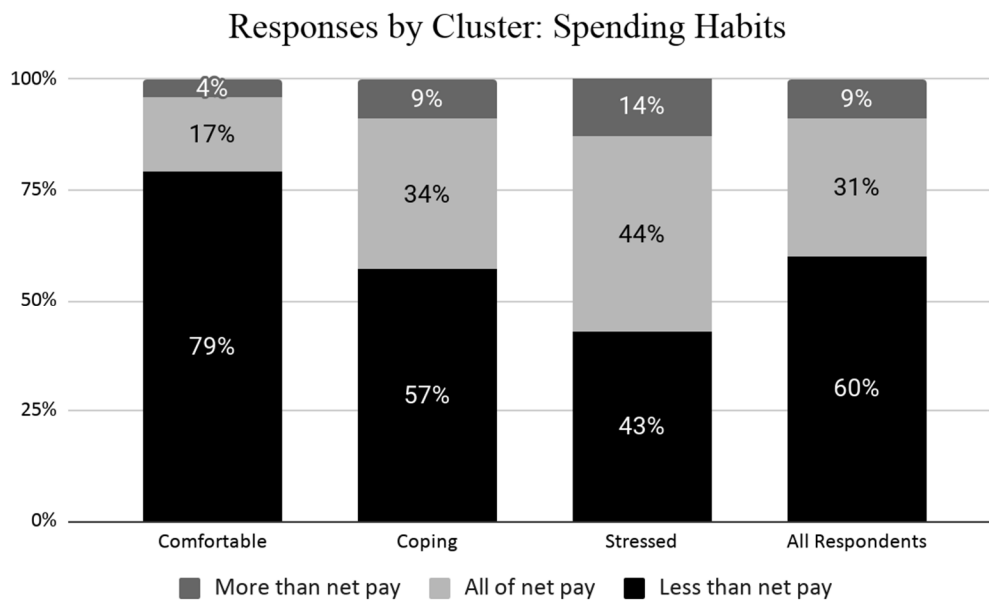


Figure 12. This figure illustrates between-cluster variation in responses to the question “in a typical or average pay period, which of the following tends to be true for you”. Available responses are “I spend more than my net pay”, “I spend all of my net pay” and “I spend less than my net pay.” For example, 79% of the comfortable group selected “less than”, 17% selected “all of” and 4% selected “more than”.

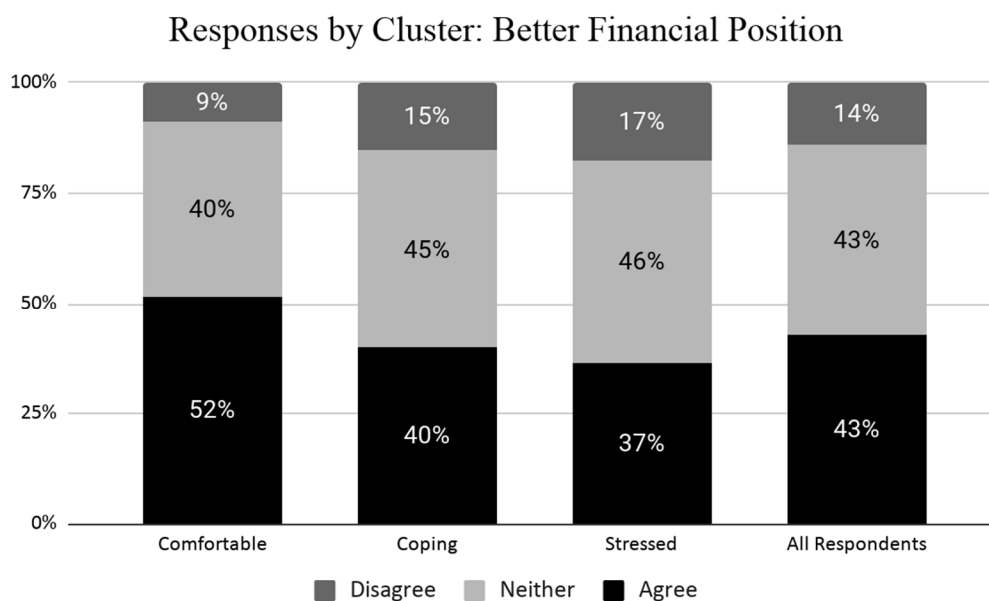


Figure 13. This figure illustrates between-cluster variation in responses to the question that asked respondents how strongly they agree or disagree (on a 7-point Likert scale) with the statement “I am in a better financial position than I was a year ago”. For example, 52% of the comfortable group selected “strongly agree” or “agree”, 40% selected “somewhat agree”, “neither” or “somewhat disagree” and 9% selected “disagree” or “strongly disagree”.

Responses to the non-core questions also reveal three additional insights, which do not appear to have been documented in the empirical literature on financial well-being. First (and unsurprisingly), financial stress appears to have a direct impact on personal well-being. In particular, and relative to the comfortable group, the stressed group is twice as likely to report that stress related to personal finances has impacted their workplace performance⁶ and more than three times as likely to report that they feel overwhelmed by their debt⁷. Second, there does not appear to be any relationship between budgeting behaviours and financial well-being. In particular there does not appear to be any difference in the frequency with which the groups budget their expenditures, or the tools that they use to construct those budgets. The comfortable group does report slightly greater success at following their budget, but not to a considerable degree. These findings are consistent with the results presented with Lee et al. (2020), where a “propensity to plan” scale, half of whose questions involve budgeting, does not appear to explain any of the observed variation in FWB scores in the US population. Third, the stressed and comfortable groups have different motivations for saving, and different ideas on how to improve their financial position. In particular, among respondents that reported they were trying to save more than they had in the previous year, stressed respondents were twice as likely to identify fear

⁶ 39% of the stressed group selected either “strongly agree” or “agree” when asked how strongly they agree or disagree (on a 7-point Likert scale) with the statement “stress related to personal finances has had an impact on my workplace performance”, as compared to 20% of the comfortable group.

⁷ 29% of the stressed group selected either “strongly agree” or “agree” when asked how strongly they agree or disagree (on a 7-point Likert scale) with the statement “I feel overwhelmed by my debt”, as compared to 8% of the comfortable group.

of uncertainty as their main motivation for increased savings, and considerably less likely to identify an improved financial situation⁸. It is also interesting to note that the stressed group is more likely to select “earn more” as compared to “spend less”, when asked to identify the single most effective way to improve their financial situation⁹ (the opposite is true of the comfortable group).

4.4. Representative members

Recall that in the process of arranging respondents into clusters, the PAM algorithm identifies the central member, or medoid, of each group. We may rigorously interpret the medoid as the “most typical” or “most representative” member of the cluster, in the sense that the average distance between this member and all other members of the cluster is at (or very close to) a minimum. In other words, to be a member of a given cluster is to have provided responses to the core survey questions that are very similar to those provided by the cluster’s medoid. The table below summarizes the responses of each cluster’s medoid to select core questions, in 2018.

As expected, the most typical member of the stressed group saves a small fraction of his pay, would find it difficult to deal with a brief financial setback, and places more importance on high wages than any other aspect of employment. Though the stressed medoid does have the lowest household income of the three, it is very interesting to note that he is in the adjacent income bracket to the comfortable medoid; this provides a concrete example of income playing a limited role in determining financial well-being.

Unsurprisingly, the comfortable medoid saves the highest proportion of his paycheck, is best placed to deal with a brief financial setback, and places more emphasis on work/life balance than any other aspect of his employment (such as higher wages). It is very telling that the comfortable medoid is not the highest earning medoid, indeed he earns less than half of what the coping method does.

The coping medoid lies somewhere between the other two in terms of financial resilience. Her savings habits are identical to those of the stressed medoid but her employment priorities are identical to those of the comfortable medoid; in other words, she has things in common with both. Interestingly, and as noted previously, the comfortable medoid earns nearly twice as much as the coping medoid.

5. Summary and conclusions

This paper applies cluster analysis to eleven years’ worth of responses to the CPA survey. Both the data set and statistical methodology are new to the empirical literature on financial well-being. Because of its longevity and consistency, the CPA data offers unique insights into the financial circumstances of employed Canadians, over a very long window. And our use of cluster analysis

⁸ Respondents that reported they were trying to save more than they had in the previous year, were asked to identify their main motivation for trying to save more. Available responses were “fear of uncertainty”, “major purchase”, “nearing retirement” and “in a better financial position”. 24% of stressed respondents selected “fear of uncertainty”, as compared to 12% of the comfortable group. 44% of the comfortable group selected “in a better financial position”, as compared to 29% of the stressed group.

⁹ 25% of the stressed group selected “earn more” and 14% selected “spend less”. The corresponding proportions for the comfortable group were 17% and 21%, respectively.

demonstrates that it is possible to generate rigorous insights into financial well-being, without needing to measure it directly.

Our analysis demonstrates that financial well-being (among employed Canadians) exists on a spectrum, with financially stressed individuals on one end and financially comfortable individuals on the other. We find that financial stress has been widespread among employed Canadians for many years, and that hallmarks of stressed individuals are difficulty saving and low financial resilience, as opposed to simple demographics such as age or income.

Our findings supplement many findings from the nascent empirical literature on financial well-being. For instance, we are able to confirm the definite, but limited, roles played by age and income, as well as the critically important roles of savings and financial resilience, in determining financial well-being. We are also able to confirm that financial stress appears to have a direct impact on personal well-being (which is unsurprising), and that budgeting behaviours do not appear to be related to financial well-being.

That financial well-being is not simply determined by income is, in fact, encouraging. Indeed, it suggests that there is considerable scope for financial planners and counsellors to empower individuals to improve their resilience (possibly by buttressing their savings habits), regardless of how much those individuals earn. To this end, it is critical to understand how some low-income individuals are able to avoid financial stress, and why some high-income individuals succumb to it. One approach might be to ask a large sample of planners and counsellors to assess all of their clients' levels of financial resilience (e.g. by asking the "\$2,000" and "paycheque delay" questions) and identify individuals whose resilience is surprising, in light of their income. Follow-up interviews could then be conducted, with a view to understanding the factors that have modulated the individuals' incomes to such an extent (e.g. recent job loss or divorce, health status, number and nature of dependents). Analysis of the resulting transcripts, possibly through an academic study, would ostensibly allow planners or counsellors to offer advice and/or strategies that are specifically tailored to a client's specific extenuating circumstances.

From a methodological perspective, an important avenue for future research is to better understand the dynamics of (i) the number of clusters, (ii) the characteristics of each cluster and (iii) the distinctiveness of the clusters. In particular, and once enough data becomes available, it will be interesting to assess the impact of the ongoing global pandemic on these features. In addition, and as more individuals complete the survey in multiple years, it will be important to understand how and why individuals move up and down the spectrum. Finally, it would be interesting to apply supervised learning techniques (e.g. classification trees) to generate a more rigorous understanding of what drives the "soft labels" obtained in this initial exercise.

Acknowledgments

This work was made possible through generous funding from (i) the Centre for Quantitative Analysis and Modelling (CQAM) at the Fields Institute and (ii) the Canadian Payroll Association. The authors would like to thank Peter Tzanetakis (Canadian Payroll Association), Alec Milne (Framework Partners), R. Mark Reesor (Wilfrid Laurier University) and Matt Davison (Western University) for valuable input. They would also like to thank Behzad Ghafouri, Longlong Feng, Agassi Iu and Jared Arcilla for valuable technical assistance.

Conflict of interest

The authors declare no conflicts of interest in this paper.

References

- Arbelaitz O, Gurrutxaga I, Muguerza J, et al. (2013) An extensive comparative study of cluster validity indices. *Pattern Recogn* 46: 243–256.
- Bezdek J, Pal N (1998) Some new indexes of cluster validity. *IEEE T Syst Man Cy B* 28: 301–315.
- Castro-González S, Fernández-López S, Rey-Ares L, et al. (2020) The influence of attitude to money on individuals' financial well-being. *Soc Indic Res* 148: 747–764.
- Collins JM, Urban C (2020) Measuring financial well-being over the lifecourse. *Eur J Financ* 26: 341–359.
- Financial Well-being: The Goal of Financial Education. Consumer Financial Protection Bureau, 2015. Available from: <https://www.consumerfinance.gov/data-research/research-reports/financial-well-being/>.
- CFPB Financial Well-Being Scale: scale development and technical report. Consumer Financial Protection Bureau, 2017. Available from: <https://www.consumerfinance.gov/data-research/research-reports/financial-well-being-technical-report/>.
- Cude BJ, Danna D, Kabaci MJ (2016) Financial knowledge and financial education of college students, In: Xiao JJ, *Handbook of Consumer Finance Research*, Eds., Springer.
- DeVaney S (2016) Financial issues of older adults, In: Xiao JJ, *Handbook of Consumer Finance Research*, Eds., Springer.
- Foss AH, Markatou M, Ray B (2018) Distance metrics and clustering methods of mixed-type data. *Int Stat Rev* 87: 80–109.
- Gower JC (1971) A general coefficient of similarity and some of its properties. *Biometrics* 27: 857.
- Gutter M, Copur Z (2011) Financial behaviors and financial well-being of college students: Evidence from a national survey. *J Fam Econ Issues* 32: 699–714.
- Hinton G, Sejnowski TJ (1999) *Unsupervised learning: Foundations of neural computation*. Massachusetts Institute of Technology.
- Iramani R, Lutfi L (2021) An integrated model of financial well-being: The role of financial behaviour. *Accounting* 7: 691–700.
- Kaufman L, Rousseeuw PJ (1987) Clustering by means of medoids, In: Dodge Y, *Statistical Data Analysis Based on the L1 Norm and Related Methods*, Eds., Amsterdam: North Holland/Elsevier, 405–416.
- Kempson E, Finney A, Poppe C (2017) Financial well-being: A conceptual model and preliminary analysis. *Project Note #3–2017, Consumption Research Norway-SIFO, Oslo and Akershus University College of Applied Sciences*.
- Kempson E, Poppe C (2018) Understanding financial well-being capability: A revised model and comprehensive analysis. *Professional Note #3–2018, Consumption Research Norway-SIFO, Oslo and Akershus University College of Applied Sciences*.
- Lee JM, Lee J, Kim KT (2020) Consumer financial well-being: Knowledge is not enough. *J Fam Econ Issues* 41: 218–228.
- Likas A, Vlassis N, Verbeek JJ (2003) The global k-means clustering algorithm. *Pattern Recogn* 36: 451–461.

- Lusardi A, Schneider D, Tufano P (2011) Financially fragile households: Evidence and implications, In: *Brookings Papers on Economic Activity*, Eds., Spring, 83–134.
- Madhulatha TS (2011) Comparison between k-Means and k-medoids clustering algorithms. *Adv Comput Inform Technol Commun Comput Inform Sci*, 472–481.
- Moscone F, Tosetti E, Vittadini G (2016) The impact of precarious employment on mental health: The case of Italy. *Soc Sci Med* 158: 86–95.
- Nanda AP, Banerjee R (2021) Consumer's subjective financial well-being: A systematic review and research agenda. *Int J Consum Stud*.
- Nielsen RB, Fletcher CN, Bartholomae S (2016) Consumer finances of low-income individuals, In: Xiao JJ, *Handbook of Consumer Finance Research*, Eds., Springer.
- Prawitz AD, Cohart J (2016) Financial management competency, financial resources, locus of control, and financial wellness. *J Financ Counsel Plan* 27: 142–157.
- Prawitz AD, Garman ET, Sohaindo B, et al. (2006) InCharge Financial Distress/Financial Well-Being Scale: development, administration, and score interpretation. *J Financ Counsel Plan* 17: 34–50.
- Prawitz AD, Kalkowski JC, Cohart J (2012) Responses to economic pressure by low-income families: Financial distress and hopefulness. *J Fam Econ Issues* 34: 29–40.
- Reynolds AP, Richards G, Iglesia BD, et al. (2006) Clustering rules: A comparison of partitioning and hierarchical clustering algorithms. *J Math Model Algorithms* 5: 475–504.
- Rousseeuw PJ (1987) Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 20: 53–65.
- Salignac F, Marjolin A, Reeve R, et al. (2019) Conceptualizing and measuring financial resilience: A multidimensional framework. *Soc Indic Res* 145: 17–38.
- Strömback C, Skagerlund K, Västfjäll D, et al. (2020) Subjective self-control but not objective measures of executive functions predicts financial behaviour and well-being. *J Exp Financ* 27.
- Vlaev I, Elliot A (2014) Financial well-being components. *Soc Indic Res* 118: 1103–1123.
- Xiao JJ (2016) Consumer financial capability and wellbeing, In: Xiao JJ, *Handbook of Consumer Finance Research*, Eds., Springer.
- Xu D, Tian Y (2015) A comprehensive survey of clustering algorithms. *Ann Data Sci* 2: 165–193.



AIMS Press

© 2021 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)