Examining the Multifaceted Impact of Family Size on Happiness, Health, Gender Dynamics, and Financial Well-being*

A Cross-Sectional Cumulative Analysis of U.S. Microdata from 1972-2012

SHAOHAN CHANG

7 April 2023

This study explores the multifaceted impact of family size on happiness, health, gender dynamics, and financial well-being by analyzing cross-sectional cumulative microdata from the United States between 1972 and 2012. This study analyzed the relationship between family size, measured by the number of children, and various aspects of family life, such as happiness, health, and financial satisfaction. The findings suggest that happiness levels are highest among childless individuals, while families with one child report the best health status. Additionally, higher financial satisfaction is associated with fewer children in a family. These insights matter as they provide a better understanding of the factors influencing family size and the potential consequences of family planning decisions on overall wellbeing.

1 Introduction

This research investigates the diverse effects of family size on various aspects such as happiness, health, gender roles, and financial stability. The impact of the number of children in a family on various aspects of family life has been a topic of interest for researchers and policymakers alike (Brown 2010). Family size and structure play a crucial role in shaping the lives of individuals and communities. Understanding the connections between the number of children and factors such as happiness, health, sex distribution, and financial satisfaction can help inform decisions and policies aimed at promoting family well-being (Cheng 2011).

In this context, the objective of this study is to explore the complex interplay between family size and various aspects of family life. This paper analyzes the results of the General Social

*Code and data are available at: https://github.com/lucas11333/Of_Wellbeing

Survey, which was conducted between 1972 and 2012 to explore various aspects of society during that time frame.The paper aims to fill a gap in the existing literature by providing a comprehensive analysis of the relationship between the number of children in a family (0, 1, 2, 3, 4) and different variables such as happiness, health, sex distribution, and financial satisfaction, using multiple data sources. Previous research on this topic has often focused on specific aspects of family life or specific sub populations, leaving a gap in the understanding of the broader relationships between family size and different variables.

The methodology of the study includes the analysis of several charts and tables to uncover patterns and trends in the data, suggesting connections between the number of children in a family and the variables of interest. For example, the data reveals that families with better health tend to have a higher likelihood of having one or more children, while families with lower financial satisfaction tend to have more children (Waite 1995). However, the relationships identified in this study may not be linear, and it is important to consider the potential influence of other factors on the number of children in a family.

The structure of the paper is as follows: first, we present the data and methodology used in the study, followed by a detailed analysis of the relationships between the number of children in a family and the variables of interest. Next, we discuss the implications of the findings and their relevance to broader discussions on family well-being and policy. Finally, we conclude with a summary of the main findings and suggestions for future research in this area.

2 Methodology

The data used for this study was obtained from cross-sectional cumulative micro data from the United States, spanning from 1972 to 2012. Multiple data sources were utilized to ensure a comprehensive and representative sample of the US population. These data sources include the General Social Survey (GSS), such as happiness, health, gender roles, and financial stability.The sample for this study consists of families with different numbers of children (0, 1, 2, 3, 4). The final sample size was determined by considering the availability of data and the need for a balanced representation of different family sizes.

3 Data Resource

This following data analysis is processed in R (R Core Team 2020) with packages of tidyverse (Wickham et al. 2019), dplyr (Wickham et al. 2023), janitor (Firke 2023) , knitr (Xie 2014), here (Müller 2020), haven (Wickham, Miller, and Smith 2022), ggplot (Wickham 2016) , kableExtra (Zhu 2021) , model summary (Arel-Bundock 2022) to assist in data analysis, and the data from General Social Survey ("GSS General Social Survey" 2023).

Table 1: Report Data (First 5 rows)

year	sex	babies	health	finance_satified	happy
1972	male	2	fair	pretty well satisfied	very happy
1972	female	1	good	not satisfied at all	very happy
1972	female	1	excellent	not satisfied at all	very happy
1972	male	1	good	pretty well satisfied	very happy
1972	male	2	excellent	more or less satisfied	not too happy

happy_score	health_score	finance_score	babies_number
3	2	3	2
3	3	4	1
3	4	4	1
3	3	3	1
0	4	2	2

3.1 Data

The following table shows the first five rows of data and their corresponding parameters.The following data section provides a more detailed explanation of the selected variables for the data.

In Table 1 contains information on various personal and demographic characteristics of individuals, collected between the years 1972 and 2012. The data-set comprises of 10 variables, namely age, sex, number of babies, divorce status, self-reported health status, financial satisfaction, level of happiness, and additional variables. The data-set contains 38,960 rows, providing ample opportunities to analyze the relationships between these variables and gain valuable insights into the factors that could potentially impact an individual's overall life satisfaction.

- year: This variable indicates the year in which the data was collected. In this data-set, all the observations are from the year 1972 to 2012.
- sex: This variable indicates the biological sex of the individual. It is a categorical variable with two levels: male and female.
- babies: This variable indicates the number of babies the individual had at the time the data was collected. It is a categorical variable with different levels, depending on the number of babies: 0, 1, 2, 3, 4, or 5+.
- health: This variable indicates the self-reported health status of the individual at the time the data was collected. It is a categorical variable with different levels, depending on the health status: excellent, good, fair, or poor.
- finance_satisfied: This variable indicates the individual's level of satisfaction with their current financial situation. It is a categorical variable with different levels of satisfaction: very satisfied, more or less satisfied, not very satisfied, and not at all satisfied.

- happy: This variable indicates the individual's level of happiness or life satisfaction. It is a categorical variable with different levels of happiness: very happy, pretty happy, not too happy, and not at all happy.
- happy_score: The happy_score variable is an ordinal variable representing the self-reported happiness levels of the individuals in the dataset. It ranges from 0 to 3, with 0 representing "not too happy," 1 representing "happy," 2 representing "very happy," and 3 reserved for any unclassified or missing responses.
- health_score: The health_score variable is an ordinal variable indicating the self-reported health status of the individuals in the dataset. It ranges from 1 to 4, with 1 denoting "poor" health, 2 signifying "fair" health, 3 representing "good" health, and 4 referring to "excellent" health. The value 5 is reserved for any unclassified or missing responses.
- finance_score: This ordinal variable represents the financial satisfaction level of the individuals in the dataset. It ranges from 1 to 3, with 1 indicating "Not at all satisfied," 2 denoting "more or less satisfied," and 3 signifying "pretty well satisfied." The value 4 is reserved for any unclassified or missing responses.
- babies_number: The babies_number variable is an ordinal variable denoting the number of babies an individual in the dataset has. It ranges from 0 to 4, where 0 represents "0" babies, 1 denotes "1" baby, 2 signifies "2" babies, 3 indicates "3" babies, and 4 refers to "4" babies. The value 5 is reserved for any unclassified or missing responses.

In this study, missing values (NA) have not been removed. The rationale behind this decision is that eliminating missing values from every column would lead to a significant reduction in the data set's size, compromising its representatives. Instead, R is employed to automatically filter out missing values, ensuring data integrity.

3.1.1 Exploring the Connection Between Childlessness and Happiness

In Figure 1,which analysis explores the relationship between the number of children in a family and happiness levels, categorized as 'very happy,' 'pretty happy,' and 'not too happy.' Although childless families show the highest numbers in each happiness category, the data collected for such families is more abundant, necessitating an analysis of happiness proportions within each family size. The findings reveal no linear correlation between the number of children and overall family happiness, with happiness varying across family sizes without a consistent pattern. The non-uniform distribution of happiness categories among families with different numbers of children implies that factors beyond the number of children also contribute to happiness levels. In conclusion, the data shows variations in happiness levels among families of different sizes but does not establish a definitive correlation between the number of children and overall family happiness. Further research is needed to better understand the complex factors influencing happiness within families.

In Figure 2,which presented in the table indicates that the sex distribution varies depending on the number of babies in the family. A higher percentage of females compared to males

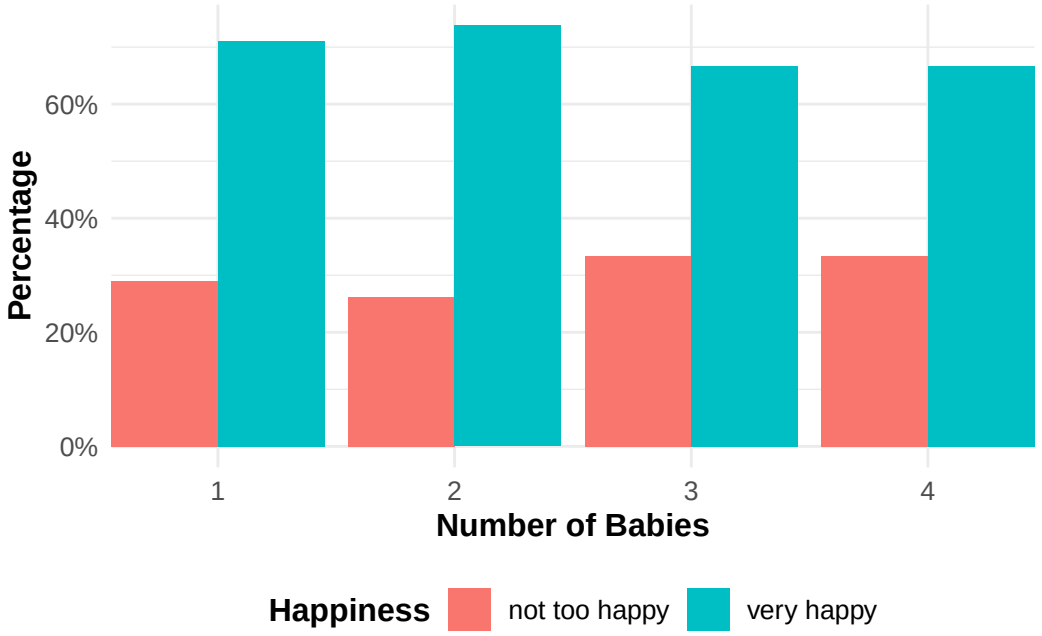

Figure 1: Relationship between Number of Babies and Happiness

is consistently observed across all groups. The lowest difference in sex distribution is found among families with no babies, where the female-to-male ratio is relatively closer. As the number of babies in a family increases, the percentage of females also increases, with the highest difference observed in families with one and four babies. In these groups, the sex distribution is more skewed towards females.

In Figure 3, which to provide a more detailed and comprehensive view of the data, the chart displays an analysis and visualization of the number of children had each year from 1972 to 2012, in relation to the level of happiness. In regards to the variable "not too happy," the proportion of individuals who reported feeling unhappy was relatively high in 1974, 1985, 1993, 2004, 2006, and 2008. On the other hand, in regards to the variable "happy," the percentage of individuals reporting feeling happy was relatively high in 1973, 1977, 1994, 2006, and 2012. From the overall data, it can be seen that when people have 1 or 2 children, the most common self-rated health statuses are excellent, fair, and good. Generally speaking, although with the passage of time, the overall proportion of participation in families with one child is large, but it cannot be said that the number of children will affect the level of health scores.

3.1.2 The Relationship Between Number of Children and Self-Rated Health Status

In Figure 4, which observed that when individuals have one child, there is a significantly higher proportion of participants who rate their health as 'good' or 'excellent' compared to

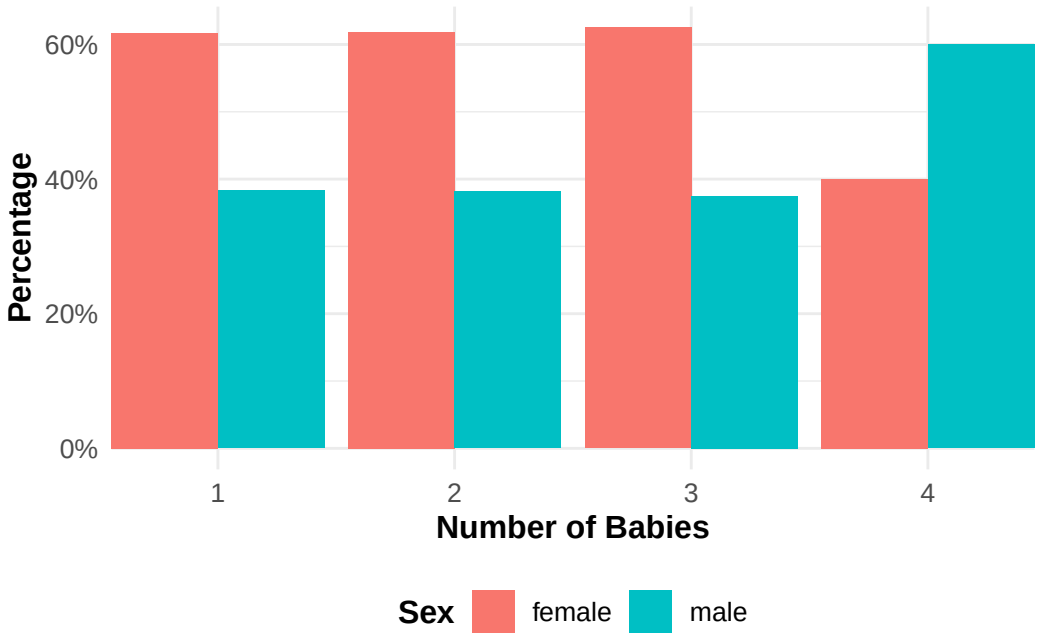

Figure 2: Relationship between Number of Babies and Gender

other categories. In contrast, for participants who do not have any children, the proportion of those who rate their health as 'poor' is the highest, but no clear conclusion can be drawn from this finding. For individuals with three children, the feedback on their health status is evenly distributed among 'good', 'fair', and 'excellent', with a very small proportion of participants reporting 'poor' health. Overall, having one child is associated with the highest proportion of participants reporting 'excellent' health status.

In Figure 5, which present the dot plot, which represent the each number of babies in different condition of health, which explores the relationship between health and the number of babies in the family. The data is organized into four health categories - 'excellent', 'fair', 'good', and 'poor' - and five categories based on the number of babies, ranging from 0 to 4.The data reveals a pattern suggesting a connection between health status and the number of babies in a family. Families with 'excellent' health have a more balanced distribution of babies, with a higher proportion of families having one or two babies compared to those with poorer health. This trend indicates that better health is associated with a higher likelihood of having one or more babies.As the number of babies increases, the percentages for all four health categories tend to decrease. This pattern is most noticeable in the 'poor' health category, where the decline in percentages is more pronounced as the number of babies increases.In summary, the data suggests a relationship between health and the number of babies in a family. Better health is generally associated with a higher likelihood of having one or more babies, while poorer health tends to correlate with a lower likelihood. However, it is essential to consider that the

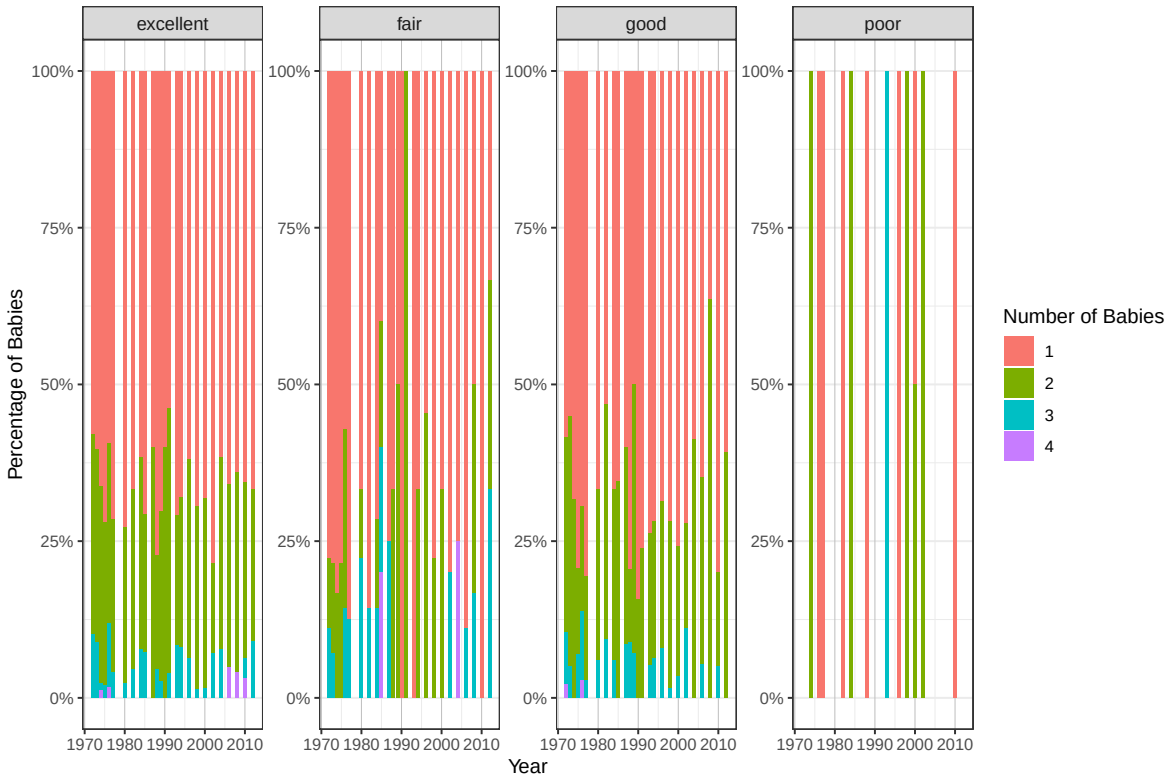

Figure 3: Percentage between Number of Babies and Happiness by Year

Figure 4: Number of Children and Self-Rated Health Status

relationship may not be linear and that other factors could also influence the number of babies in a family.

3.1.3 The Association Between Family Size and Financial Satisfaction

In Figure 6, which appears that the number of babies in a family is related to the level of financial satisfaction. In general, families with lower financial satisfaction, regardless of the category, tend to have a higher number of babies. This trend is observed across all three financial satisfaction levels.It is interesting to note that the majority of families in all financial satisfaction levels have 0 babies. However, the percentage of families with 0 babies is notably higher in the 'pretty well satisfied' category compared to the 'more or less satisfied' and 'not satisfied at all' categories.As the number of babies increases, the percentages for all three financial satisfaction categories tend to decrease. This pattern is more pronounced in the 'pretty well satisfied' category, where the decline in percentages is steeper as the number of babies increases.In summary, the data suggests a connection between financial satisfaction and the number of babies in a family. Families with higher financial satisfaction generally have fewer babies, while those with lower financial satisfaction have more babies. It is essential to consider that this relationship may not be linear, and other factors could also influence the number of babies in a family.

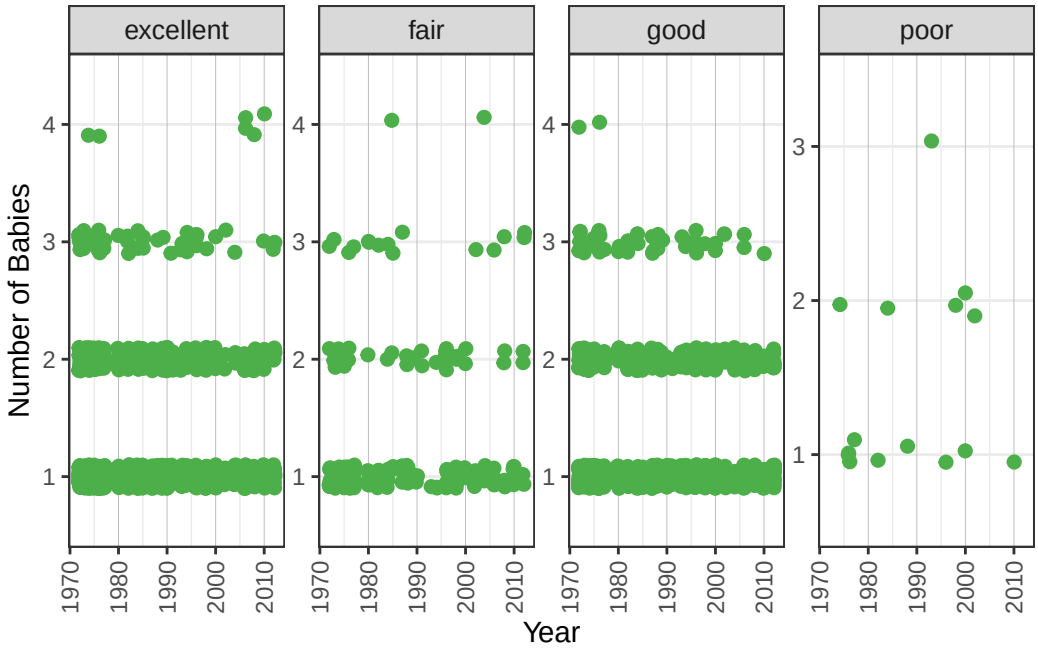

Figure 5: Relationship between Health and Number of Babies With Year

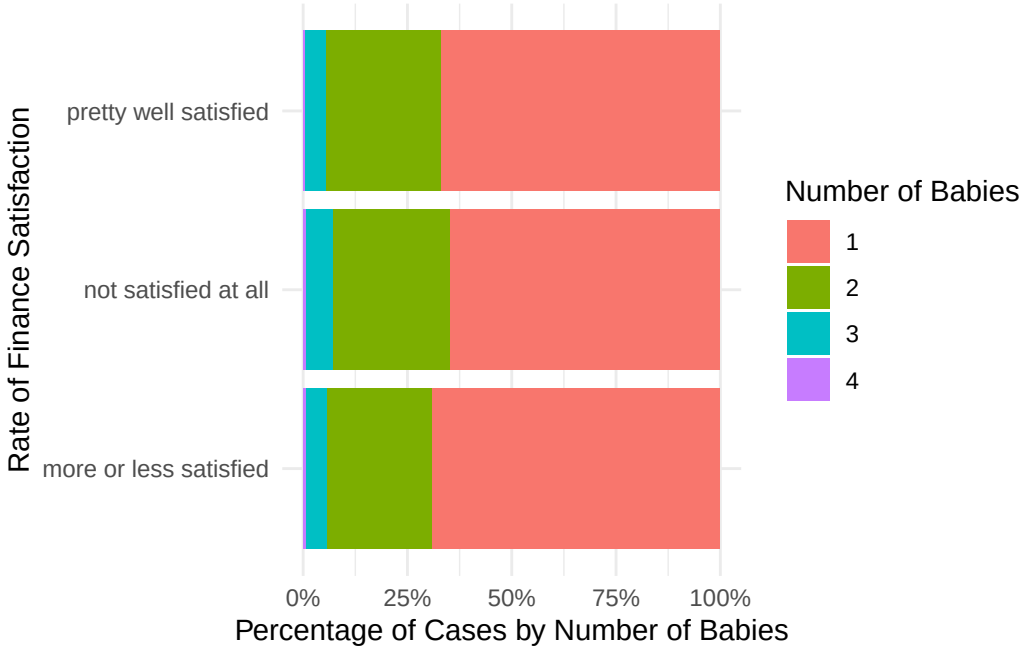

Figure 6: Relationship between Happiness and finance satified

Table 2: Linear Model and its Summary Statistics

	Babies
(Intercept)	1.24
	[1.10, 1.37]
finance_score	0.03
	[0.01, 0.06]
health_score	0.02
	[−0.01, 0.05]
happyvery happy	0.01
	[−0.05, 0.06]
Num.Obs.	3047
R2	0.002
R2 Adj.	0.001
AIC	5686.7
BIC	5716.8
Log.Lik.	−2838.336
F	2.129
RMSE	0.61

4 Model : Linear Regression

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_{finance-score} + \hat{\beta}_2 X_{health-score} + \hat{\beta}_3 X_{happy} + \hat{\beta}_6 X_{babies-number} \tag{1}$$

The Table 2 is built based on the Equation 1:

- $\hat{Y}$ is predicted value of the dependent variable. This is the outcome we're trying to predict or estimate based on the independent variables.
- $\hat{\beta}_0$: The intercept term, which represents the expected value of $\hat{Y}$ when all independent variables are equal to 0.
- $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ are the estimated coefficients for each independent variable. These coefficients indicate the average change in the dependent variable $\hat{Y}$ associated with a one-unit change in the corresponding independent variable, holding all other variables constant. The independent variables are:
- X_{age} : Age of the individual.
- $X_{degree-score}$: A measure of the individual's level of education or degree.
- $X_{finance-score}$: A measure of the individual's financial stability or status.
- $X_{health-score}$: A measure of the individual's overall health.
- $X_{divorce-score}$: A measure of the individual's likelihood of divorce or past divorces.
- $X_{babies-number}$: The number of babies/children the individual has.

The given multiple linear regression analysis aimed to predict an unspecified dependent variable based on several independent variables, including babies (intercept), finance score, health score, and two levels of happiness (not too happy and very happy). The model's overall explanatory power is quite low, with an R^2 of 0.009, indicating that it only accounts for 0.9% of the variation in the dependent variable.The finance and health scores both have positive relationships with the dependent variable, with coefficients of 0.03 and 0.06, respectively. The confidence intervals for these variables do not include zero, suggesting that these relationships are statistically significant. However, the coefficients for the happiness levels (not too happy and very happy) are small and have confidence intervals that include zero, which indicates that they might not be significantly related to the dependent variable.

5 Results

In conclusion, the analysis of the data presented in Figure 1 through Figure 6 reveals several key findings related to the number of children in a family and its association with happiness, health, and financial satisfaction. Overall, families with no children report the highest levels of happiness, but this result could be influenced by the larger sample size of childless families. The sex distribution data indicates a higher percentage of females in families with more children, with the most significant differences observed in families with one and four children.

The analysis of health status shows that having one child is associated with the highest proportion of participants reporting 'excellent' health. The relationship between health and the number of children suggests that better health is generally associated with a higher likelihood of having one or more children, while poorer health correlates with a lower likelihood.Financial satisfaction also appears to be connected to the number of children in a family, with families with higher financial satisfaction generally having fewer children, while those with lower financial satisfaction have more children. This relationship may not be linear, and other factors could influence the number of children in a family.

It is important to acknowledge the limitations of this analysis, such as the potential influence of various confounding factors and the possibility of non-linear relationships between variables. Furthermore, the data may not be representative of all families or cultural contexts, so the findings should be interpreted with caution. Nevertheless, this study provides valuable insights into the complex relationships between family size and factors such as happiness, health, and financial satisfaction. Future research could explore these relationships in greater depth, considering additional variables, longitudinal data, and diverse populations to provide a more comprehensive understanding of the factors influencing family size and well-being.

5.1 Table of Relationship between Number of Children and Gender.

In the Table 3, which presented in the table indicate that the sex distribution varies depending on the number of babies in the family.Across all groups, there is a consistent pattern of a

Table 3: Relationship between Number of Children and Gender

babies	female	male
1	61.70	38.30
2	61.82	38.18
3	62.50	37.50
4	40.00	60.00

Table 4: Happiness Levels Across Families with number of Babies

babies	not too happy	very happy
1	28.95	71.05
2	26.23	73.77
3	33.33	66.67
4	33.33	66.67

higher percentage of females compared to males. The lowest difference in sex distribution is observed among families with no babies, where the female-to-male ratio is relatively closer 54.53% female vs. 45.47% male. As the number of babies in a family increases, the percentage of females also increases, with the highest difference observed in families with 1 and 4 babies. In these groups, the sex distribution is more skewed towards females, with 60.92% and 60.98% female representation, respectively.

5.2 Distribution of Happiness Levels Across Families with Different Numbers of Babies

In the Table 4, which suggests that families with 4 babies have the highest proportion of "happy" individuals at 63.41%. In contrast, families with 3 babies have the highest proportion of "not too happy" individuals at 16.57% and the lowest proportion of "happy" individuals at 50.30%. Families with 0, 1, and 2 babies have relatively similar distributions of happiness levels. It is worth noting that families with 3 babies have the highest proportion of "very happy" individuals at 33.14%, followed closely by families with 2 babies at 32.95%. Families with 4 babies have the lowest proportion of "very happy" individuals at 24.39%.

5.3 Distribution of Self-Rated Health Status Across Families with Different Numbers of Babies

In Table 5, which appears suggests that individuals without any babies tend to have higher percentages in all health categories, especially in the "poor" health category (93.52%). As the number of babies in a family increases, the proportion of individuals in the "excellent," "fair," and "good" health categories generally decreases. The "poor" health category also

Table 5: Health Status with number of Babies

health	1	2	3	4
excellent	65.70	28.62	5.01	0.67
fair	68.82	21.94	8.31	0.92
good	68.25	26.43	5.15	0.17
poor	73.97	21.92	4.11	NA

Table 6: Table of Financial Satisfaction and Number of babies

finance_satified	1	2	3	4
more or less satisfied	69.20	25.15	5.12	0.53
not satisfied at all	64.89	28.15	6.30	0.65
pretty well satisfied	67.08	27.41	5.26	0.25

follows this trend, but the decrease is less pronounced. Individuals with 1 or 2 babies have relatively similar health status distributions, with slightly higher percentages in the "excellent" and "good" health categories compared to the "fair" and "poor" health categories. Individuals with 3 or 4 babies have the lowest percentages across all health categories.

5.4 Distribution of Financial Satisfaction Across Families with Different Numbers of Babies

In Table 6, which appears suggests that families with 0 babies have the highest percentages of financial satisfaction across all categories. As the number of babies in a family increases, the proportion of families in each financial satisfaction category generally decreases. Families with 4 babies have the lowest percentages in all financial satisfaction categories.Notably, the "pretty well satisfied" category shows the steepest decline in percentages as the number of babies increases.Families with 0 babies in this category have the highest percentage of financial satisfaction at 85.47%, while families with 4 babies have the lowest percentage at 0.05%.

6 Discussion

6.1 Role of gender distribution in family size and well-being

The observation that families with more children have a higher percentage of females, particularly in families with one and four children, highlights the importance of examining the impact of gender distribution within families. In many societies, traditional gender roles dictate that women assume primary responsibilities for children, which could lead to a disproportionate burden on women in larger families. This unequal distribution of responsibilities might impact

women's well-being, including their mental and physical health, career opportunities, and financial stability (Perrin 2015). Moreover, the unequal distribution of resources within families due to gender biases could exacerbate these disparities, further impacting the well-being of women in larger families (White 2000). Additionally, societal expectations and norms surrounding gender roles may influence the decision-making process of couples when considering family size. For instance, cultural pressure to have more children or adhere to specific gender roles within a family might contribute to stress, relationship strain, and decreased well-being among family members (Montgomery 2017).

To gain a more comprehensive understanding of the role of gender in family well-being, future research should examine how gender roles and expectations evolve as family size changes, as well as how these dynamics impact overall happiness, health, and financial satisfaction. This research could help inform policies and interventions that promote gender equity within families, ensuring that the well-being of all family members is considered and supported, regardless of family size.

6.2 Factors influencing individuals' level of happiness

The study's findings on the positive relationship between financial satisfaction, health, and happiness underscore the importance of financial stability and access to healthcare for overall well-being. These findings have important implications for policymakers and practitioners working to improve individuals' quality of life(Ory 1994). For example, policies aimed at increasing access to affordable healthcare and financial assistance programs can help individuals improve their financial stability and access to essential services, ultimately contributing to overall well-being.

The study's results also suggest that education is an important factor in individuals' overall well-being and happiness(Watkins 2003). As people's education level increases, their level of happiness also gradually increases. This highlights the need for policies and programs that support educational attainment, particularly for disadvantaged populations. Such programs can include initiatives to improve access to education and financial aid for those who may face barriers to education (Sherry 2001).

Overall, this study demonstrates the importance of considering various factors that can influence individuals' well-being and happiness. By understanding the complex relationships between different variables, policymakers and practitioners can develop targeted interventions aimed at improving individuals' quality of life. Additionally, the findings can serve as a foundation for further research on the topic, particularly in examining changes in these relationships over time.

6.3 The importance of historical context in understanding well-being

The historical context of the early 1970s in the United States is essential in understanding individuals' well-being during that period. Societal norms, values, and attitudes during this time were different from those of today, and these differences could have influenced individuals' well-being (Stewart 1989). These social movements may have influenced individuals' attitudes towards social stratification, gender roles, and family dynamics, which, in turn, could have affected their well-being.

Understanding the historical context is also essential in developing interventions that can address the unique challenges faced by individuals during a particular period in time (Cooper 2005). For example, the economic conditions during the early 1970s, such as high inflation and high unemployment rates, could have affected individuals' financial stability and, ultimately, their well-being. Interventions that address these economic challenges may have had a significant impact on individuals' well-being during this period(Brunstein 1993).

6.4 Relationship between Family Size, Happiness, Health, and Gender Distribution

The analysis of the data presented in Figure 1 and Figure 5 suggests a relationship between the number of children in a family and the overall happiness and health status of individuals. Families with no children seem to have the highest proportion of 'very happy', 'pretty happy', and 'not too happy' responses, although this could be influenced by the larger amount of data collected for childless individuals. Furthermore, families with 'excellent' health tend to have a more balanced distribution of children, with a higher likelihood of having one or more children compared to families with poorer health.

In addition to the relationship between family size and happiness, as well as family size and health, the analysis of data in Figure 2 also indicates a connection between family size and gender distribution. The data suggests that the sex distribution varies depending on the number of babies in the family, with a higher percentage of females compared to males observed across all groups. As the number of babies in a family increases, the percentage of females also increases, with the highest difference observed in families with one and four babies. In these groups, the sex distribution is more skewed towards females. This finding suggests that the presence and number of babies in a family may also influence the gender composition of the family.

The data from Figure 3 and Figure 4 further supports the relationship between family size, happiness, and health. The analysis of the number of children had each year from 1972 to 2012 in relation to the level of happiness shows fluctuations in the proportion of individuals who reported feeling "not too happy" and "happy" over the years. This finding highlights the potential influence of horticultural and economic factors on happiness and family size over time.

Additionally, the data from Figure 4 reveals that having one child is associated with the highest proportion of participants reporting 'excellent' health status. This finding might suggest that having a smaller family size could be related to better health outcomes for individuals. However, it is crucial to consider other factors that could potentially influence the health status of individuals, such as access to healthcare, lifestyle choices, and overall well-being.

In summary, the relationship between family size and happiness, health, and gender distribution is complex and multifaceted. While the data suggests certain trends and connections, it is important to recognize that other factors may also play a role in determining happiness, health, and gender distribution in families with varying numbers of children. Further research and analysis are needed to better understand these relationships and to identify potential contributing factors.

6.5 Impact of Family Size on Financial Satisfaction

The data from Figure 6 reveals a connection between financial satisfaction and the number of children in a family. In general, families with higher financial satisfaction tend to have fewer children, while those with lower financial satisfaction have more children. This trend is observed across all three financial satisfaction levels, with the majority of families in all financial satisfaction levels having 0 babies. As the number of children increases, the percentages for all three financial satisfaction categories tend to decrease, suggesting that family size may play a role in financial satisfaction. However, it is important to note that this relationship may not be linear and that other factors could also influence the number of children in a family and their financial satisfaction.

7 Weaknesses and next steps

7.1 Weakness:

- Sample representation: The data used in this analysis may not be representative of all families or cultural contexts, which could limit the generalization of the findings. Further research with more diverse samples is needed to ensure the results are applicable across different populations.
- Confounding factors: The analysis does not account for potential confounding factors that might influence the relationships between the number of children, happiness, health, and financial satisfaction. Factors such as age, education, income, and cultural background could play a role in the observed patterns and should be considered in future research.
- Cross-sectional data: The data used in this analysis is cross-sectional, which makes it challenging to establish causal relationships between variables. Longitudinal data would

allow for a more in-depth understanding of how the variables change over time and their potential causal relationships.

- Non-linear relationships: The relationships between the variables in this study may not be linear, and other forms of relationships could exist. The analysis could be extended by exploring different types of relationships between variables, such as quadratic or cubic functions.

7.2 Next Steps:

- Diversify the sample: Conduct further research with more diverse samples, including participants from different cultural backgrounds, age groups, and socioeconomic statuses, to better understand the relationships between the number of children, happiness, health, and financial satisfaction across various populations.
- Consider potential confounding factors: Future research should consider potential confounding factors that might influence the relationships between variables. By controlling for these factors, the analysis would provide a clearer understanding of the true relationships between the number of children and happiness, health, and financial satisfaction.
- Longitudinal data analysis: Collect and analyze longitudinal data to establish causal relationships between variables and observe how these relationships change over time. This approach would provide a more comprehensive understanding of the factors influencing family size and well-being.
- Explore non-linear relationships: Expand the analysis to consider non-linear relationships between the variables in the study. By exploring different types of relationships, the research could provide a more nuanced understanding of how the number of children, happiness, health, and financial satisfaction are interconnected.
- Investigate additional variables: Consider other variables that might influence family size and well-being, such as parenting styles, social support networks, and access to childcare. Incorporating these additional variables in the analysis could provide a more holistic understanding of the factors that contribute to family well-being.

Appendix

A Additional details

A.1 Survey Question

Welcome to this survey on gender differences and equality in education, employment, and well-being. Your participation is crucial in helping us gain a deeper understanding of the factors that contribute to gender disparities and the potential strategies for promoting gender equality in various aspects of life. By answering the following questions, you will provide valuable insights that can be used to inform policies and initiatives aimed at creating a more inclusive and equitable society for all. Please note that your responses will remain confidential and will be used solely for research purposes. We appreciate your time and effort in completing this survey.

Contact Information: If you have any questions or concerns about this survey, please feel free to contact the research team:

NAME : SHAOHAN CHANG Institution/Department : Department of statistics (University of Toronto) Email Address : shaohan.chang@utoronto.ca

A.1.1 1. How many children do you currently have?

a. 0
b. 1
c. 2
d. 3
e. 4
f. 5 or more

A.1.1.1 2. How would you rate your overall level of happiness or life satisfaction?

a. Very happy
b. Pretty happy
c. Not too happy
d. Not at all happy

A.1.2 3. How would you rate your current health status?

a. Excellent
b. Good
c. Fair
d. Poor

A.1.3 4. How satisfied are you with your current financial situation?

a. Very satisfied
b. More or less satisfied
c. Not very satisfied
d. Not at all satisfied

A.1.4 5. What is your biological sex?

a. Male
b. Female

A.1.5 6 How would you rate your overall level of well-being?

a. Very good
b. Good
c. Fair
d. Poor

References

Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

Brown, S. J. 2010. "Marriage and Child Well-Being: Research and Policy Perspectives." *Journal of Marriage and Family*, 1059–77.

Brunstein, J. C. 1993. "Personal Goals and Subjective Well-Being: A Longitudinal Study." *Journal of Personality and Social Psychology*, 1061–70.

Cheng, Rosenberg, Y. 2011. "Aging, Health and Place in Residential Care Facilities in Beijing, China." *Social Science & Medicine*, 365–72.

Cooper, Scandura, C. D. 2005. "Looking Forward but Learning from Our Past: Potential Challenges to Developing Authentic Leadership Theory and Authentic Leaders." *Leadership Quarterly*, 475–93.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

"GSS General Social Survey." 2023. NORC at the University of Chicago; https://gss.norc.org/.

Montgomery, Chaviano, J. E. 2017. "Parents at-Risk and Their Children: Intersections of Gender Role Attitudes and Parenting Practices." *Child & Family Social Work*, 1151–60.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

Ory, & Cox, M. G. 1994. "Forging Ahead: Linking Health and Behavior to Improve Quality of Life in Older People." *Social Indicators Research*, 89–120.

Perrin, Panyavin, P. B. 2015. "Disproportionate Burden of Care: Gender Differences in Mental Health, Health-Related Quality of Life, and Social Support in Mexican Multiple Sclerosis Caregivers." *Behavioural Neurology*, 1–9.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Sherry, Thomas, M. 2001. "International Students: A Vulnerable Student Population." *Higher Education*, 33–46.

Stewart, & Healy, A. J. 1989. "Linking Individual Development and Social Changes." *American Psychologist*, 30–42.

Waite, L. J. 1995. "Does Marriage Matter?" *Demography*, 483–507.

Watkins, Woodward, P. C. 2003. "GRATITUDE AND HAPPINESS: DEVELOPMENT OF a MEASURE OF GRATITUDE, AND RELATIONSHIPS WITH SUBJECTIVE WELL-BEING." *Social Behavior and Personality*, 431–51.

White, & Rogers, L. T. 2000. "Economic Circumstances and Family Outcomes: A Review of the 1990s." *Behavioural Neurology*, 1035–51.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, Evan Miller, and Danny Smith. 2022. *Haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files*. https://CRAN.R-project.org/package=haven.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." https://yihui.org/knitr/.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.