Does inflation make you spend less?*

Rahma Binth Mohammad

August 12, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

Table of contents

*Code and data are available at: https://github.com/rahmabinth/consumer-behaviour

1 Introduction

Consumption is the biggest indicator of a person's needs and desires. It tells us the exact amount and item people are willing to spend on. Our interest is to inquire into whether tracking consumption, along with income and prices of items through inflation, will tell us a better story about a household's financial behaviour and broader economic trends. By assessing these variables, we can uncover how inflation influences spending habits and whether an increase or decrease in the number of income earners in a region can mitigate or exacerbate the effects of rising or falling prices. The relationship between these variables can provide information in regards to improving economic stability and effective social governmental assistance to support households in managing their spending in the face of inflationary pressures. Overall, policymakers can use this information to employ strategies for the wellbeing of the state and ensure households and even states don't go deeper into debt.

Recently, inflation has become serious issue. It diminishes purchasing power, which is a problem at the individual level but also state level. On an individual level, it affects costs of living, impeding the ability to live luxuriously. On a state level, it can lead to increased costs for goods and services, reducing the competitiveness of a state's exports and making it harder for businesses to thrive on a domestic and global market. Thus, it is crucial states manage inflation as it will strength consumer and investor confidence, leading to economic stability and growth.

We used a simple linear regression model and a multiple linear regression for our analysis. For the simple linear model, we used the the change in prices over several years against the average household spending during these years. For the multiple regression model, we used the number of income earners, along with the change in the prices to predict household spending. Our estimand is the average total household expenditure including recreational usage.

Based on our model and analysis, we found that despite rising prices, households continued to spend money on necessary items, as well as recreational spending. However, recreational spending saw a slower increase over a decade. Incorporating the change in the number of income earners ensured we factored in the change in population and the constant flux of people moving to or away from the region.

This paper was designed to communicate our model and analysis in a structured manner. Section 2 will provide the datasets used for this analysis. Section 3 will explore the simple and multiple linear regression models used for this analysis, along with the justification for the use of such models. Section 4 will provide the results from that analysis. A discussion of the results and the weaknesses and biases of this type of analysis will be provided in Section 5.

2 Data

2.1 Methodology

The language and environment used for this analysis is R (R Core Team 2023), alongside the tidyverse (Wickham et al. (2019a)), rstanarm (Goodrich et al. (2022)), cansim (von Bergmann and Shkolnik 2021), ggplot2 (Wickham 2016), scales (Wickham, Pedersen, and Seidel 2023), and modelsummary (Arel-Bundock 2022) packages.

We downloaded three different datasets directly from Statistics Canada using the cansim package. We cleaned it by isolating the variables we need, removed missing values, and wrote csv files to make it easily accessible using R studio.

2.2 Sources

A good indicator of inflation is the Consumer Price Index (CPI), as it measures the change in price using a fixed basket of goods and services. Statistics Canada is a government agency that collects information on Canada's economy (Canada, n.d.) and it has the data that provides the CPI for numerous years. Statistics Canada also provides us with other datasets such as the average household expenditures in a year for all the provinces and the number of income earners in a province. We have chosen to look at Ontario solely, as we hold firmly that Ontario behaves very differently due to its high population in relation to the other provinces. The data for Canada as a whole is complicated as there are many missing values for some provinces. However, our analysis and the data provided by Statistics Canada makes it possible to reproduce our model for different provinces using the github repository for this paper. While there are other datasets available, these three datasets are provided by the same source and has data for provinces across Canada, making it easily reproducible. The expenditure and CPI data also provides the information for three additional specific categories, food, shelter, and recreation, making it easier to isolate and answer our research question. Section 2.2 will illustrate the visualised data.

This paper's objective is to address the following questions using data collected from a government census: (1) Does changes in prices affect expenditure? (2) Does changed in prices affect expenditure as the number of income earners change?

2.3 Variables and Measurement

2.3.1 Spending Data

`spending.csv`

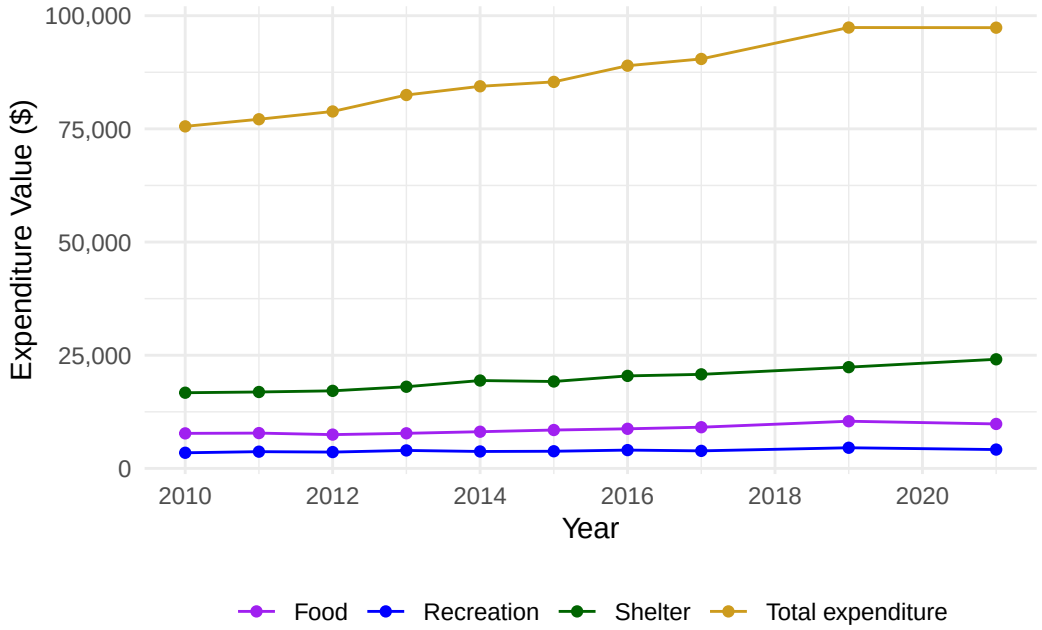

Figure 1: Household spending per year

This dataset contains the average total expenditures of the Ontario resident for the years 2010 to 2021, excluding years 2018 and 2020. It also contains the breakdown of the expenditures, food purchased (from restaurant and stores), shelter (rented and owned), utilities such as water, fuel, and electricity for the principal accommodation, clothing, transportation (private and public), healthcare, recreational goods and services.

From this dataset, we chose to select the values for all food expenditures, all shelter expenditures, all recreation expenditures, and total expenditures. This choice was made because these are principal expenditures for everyone. Recreation was added to this mix, to identify how consumers behave when inflation is high. It helped answer the question of whether consumers spend less on items that are unnecessary to survive? We define necessary items to be food, shelter, and clothing.

In Figure 1, we observe a steady increase in recreation expenditures, while food, shelter, and overall expenditures show a more rapid growth. Note for simplicity, we did not include the other categories used to calculate total expenditures in the original dataset. This is the reason the total expenditure is significantly higher than the other categories. For all four of the categories, the mean and median values were close to each other, indicating the data is not skewed and did not have many outliers. However, the standard deviation for shelter and total expenditures were higher which indicates more money was spent and thus, more variation over the years. The only variable we use for our models from this dataset is `total expenditure`. However, the model can be used for the other categories. The model is made to be easily reproducible and can be found in the github link provided.

2.3.2 Inflation Data

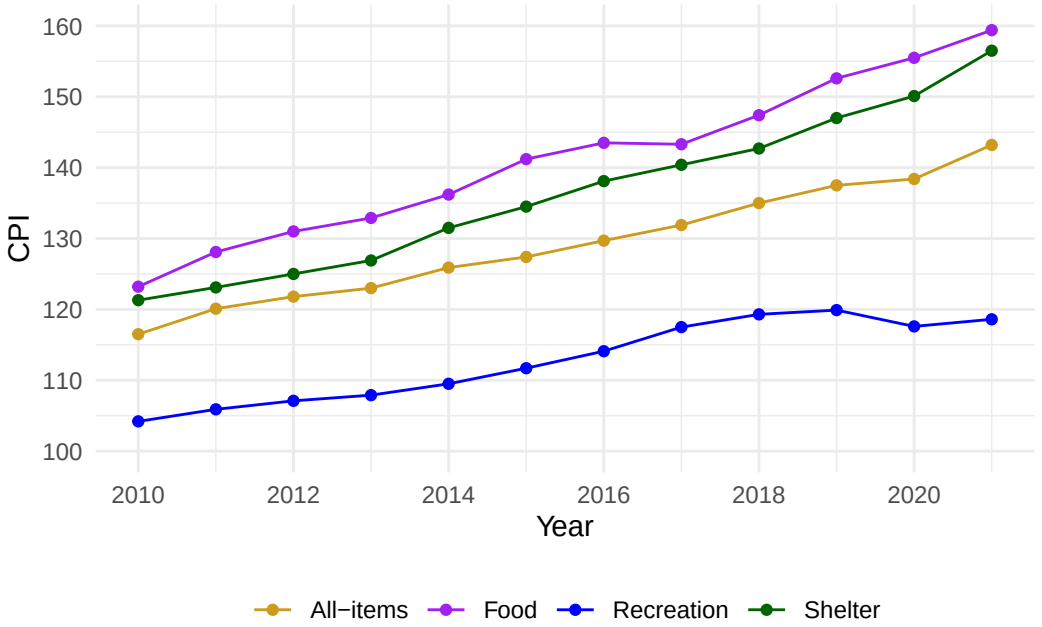

Figure 2: Consumer Price Index (CPI) per Year

`cpi.csv`

The dataset provides CPI for every year. It is not seasonally adjusted for changes as it is an annual average. Thus, it is the actual price paid by consumers and not adjusted for seasonal fluctuations. The base year for the CPI calculation is 2002 where CPI is 100. Thus, if our CPI is 120 for a year, it indicates a 20% increase since 2002. This value is useful as an indicator for inflation, as it provides the increase and comparison between years, rather than a absolute price values.

We chose to look at the annual average for CPI to keep it consistent with the annual average of expenditures. The data provides CPI for specific categories, as well as all-items. We chose to select CPI for all-items, food, shelter, and recreations, for comparison with the household spending. The prices for food and shelter increased almost 35% in the 10 years illustrated in Figure 2. However, the prices for recreation saw a decrease in 2020 and 2021, which indicates a relation to COVID-19. Our models will use the 'all-items' CPI to ensure consistency.

2.3.3 Income Data

`income.csv`

Everyone's financial situation is different, especially in Ontario. Statistics Canada provides us with this dataset that contains the average number of people in an income category. For

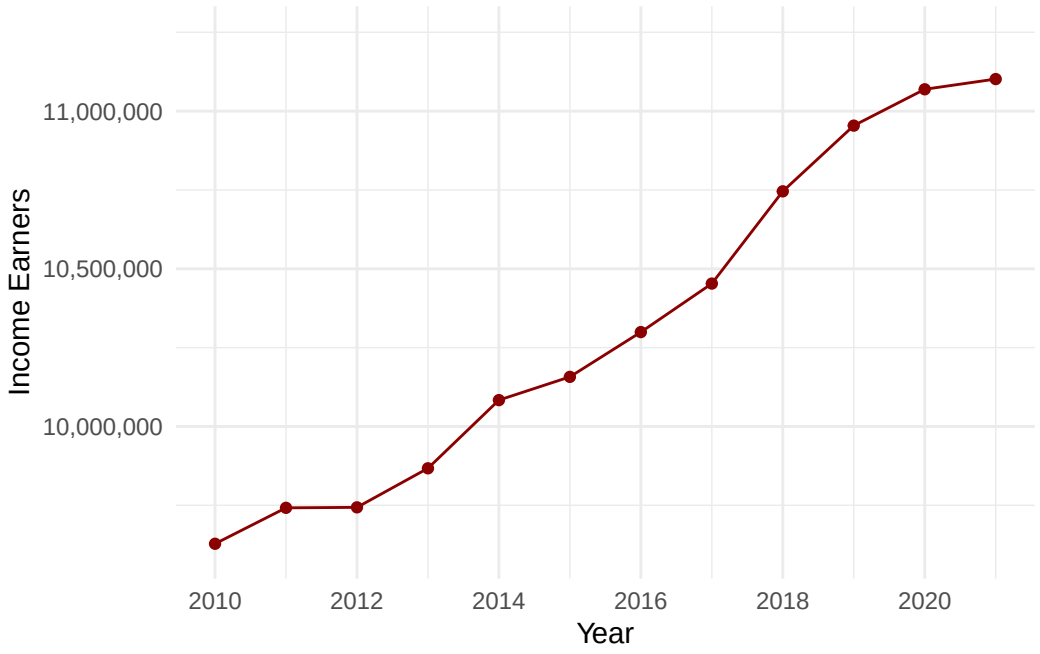

Figure 3: Number of Income Earners per Year

example, "persons with income under $5000", "persons with income of $5000 and over", "persons with income of $10 000 and over." Note for 2020, COVID-19 government income support and benefits are included in income estimates.

From this dataset, we chose to isolate the values for the category "all persons with income". This indicate all tax fillers and dependants that earn income from all sources. This was done to keep our analysis consistent, as we could not separate who was spending on what item. Our model will use the `all persons with income` variable as well. Figure 3 shows a significant increase in the number of income earners. This could be due to a population increase, specifically an increase in tax fillers and tax dependants because the data is taken directly from filed tax returns.

For our analysis, we had to use all three dataset values and thus, we chose to use data for the years 2010 to 2021, excluding 2018 and 2020, given the data was only available in all given datasets for these years. However our model does predict the behaviour in 2018 and 2020, using the other years' data. We also created numerous csv files that isolated the specific data. The variables discussed in this section became an entry in our dataset, because each variable answers a relevant question.

The variables discussed in Section 2.2 became an entry in our dataset, because each variable answers a relevant question. Household spending data was only available from years 2010-2021. This is a relatively small interval so we had limitations for our model. However, it is still useful because a lot can change in a decade and if you look at data before 2010, it will not

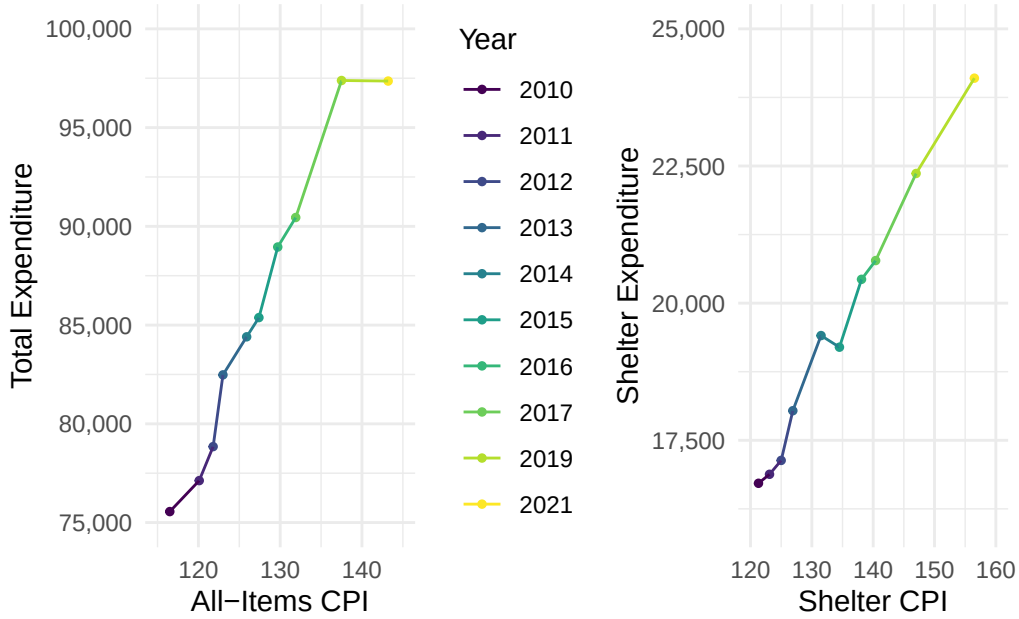

Figure 4: All-items CPI and total expenditure over the years (on the left), shelter CPI and shelter expenditure over the years (on the right)

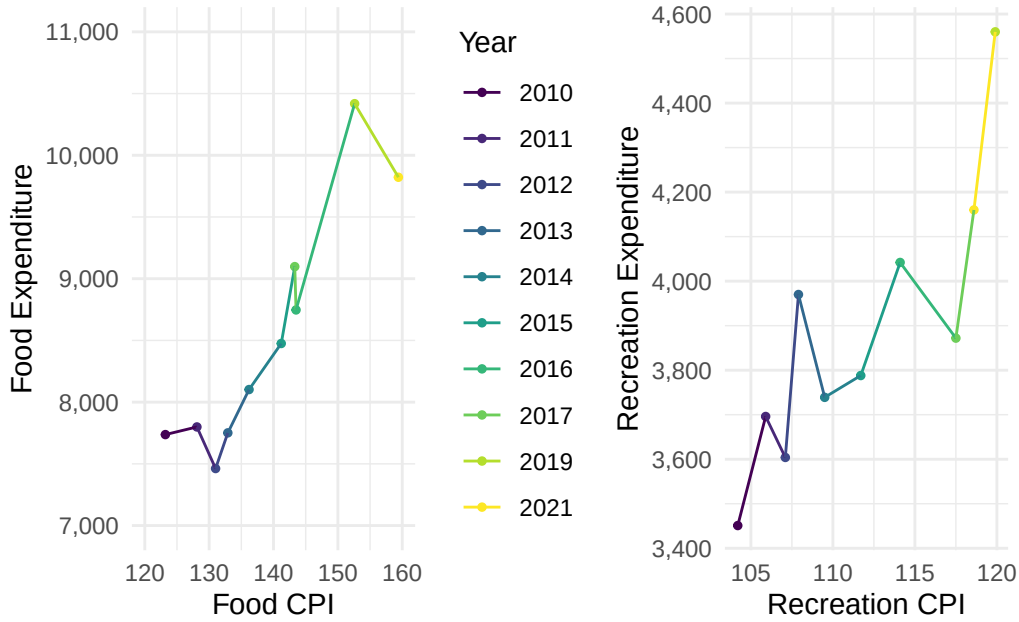

Figure 5: Food CPI and food expenditure over the years (on the left), recreation CPI and recreation expenditure over the years (on the right)

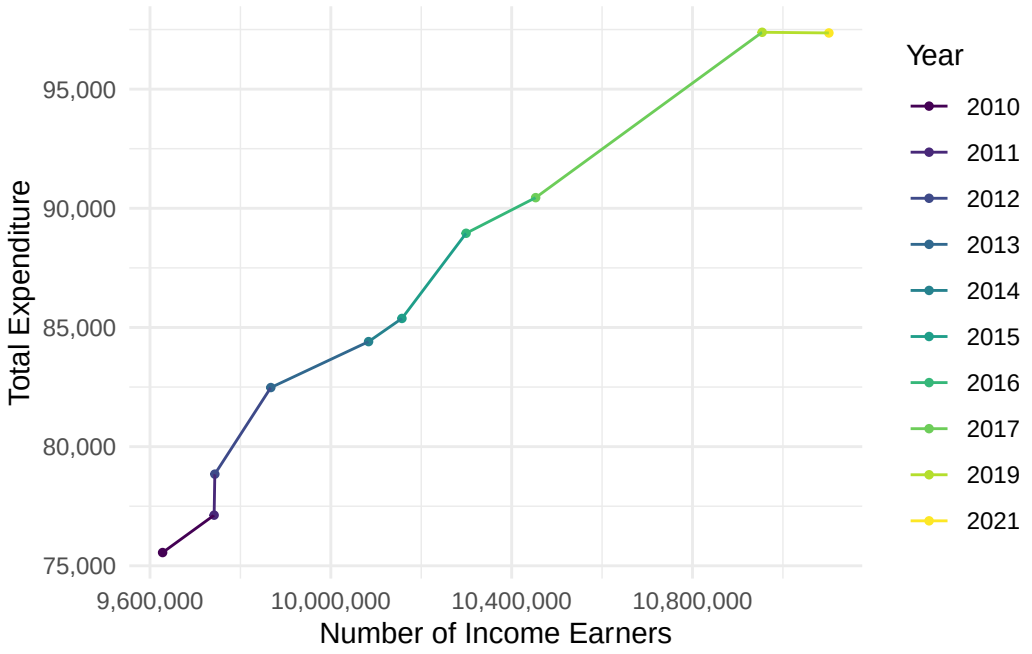

Figure 6: The number of income earners and their average expenditures

give you a clear visual of what is happening today. Thus, making it harder to make political or economical decisions that impact today's world.

Figure 4, Figure 5, and Figure 6 are all the predictor variables graphed against the total expenditure. There is an increase in spending over the years, as well as CPI and income earners. In the rare occasions, where CPI decreases, expenditure decreases as well.

3 Model

In this section, we build a model to investigate household consumer behaviour.

3.1 Model set-up

The goal of our model is to forecast the total expenditure in a household in a year, based solely on the Consumer Price Index of all-items, and the number of people in Ontario earning an income. We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022).

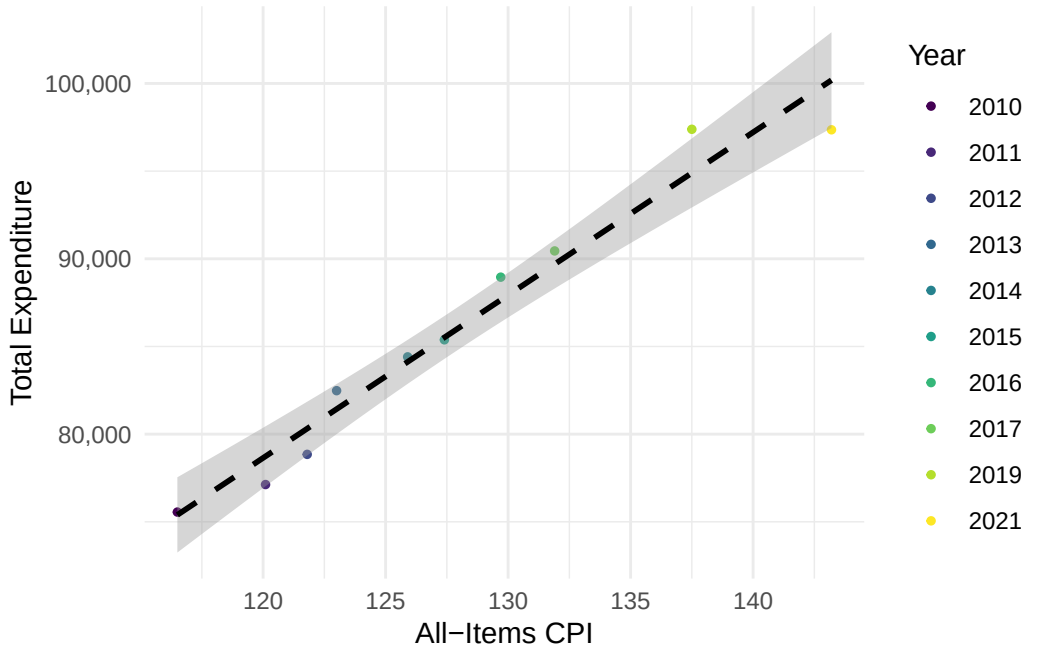

Figure 7: Relationship between CPI for all-items and total household expenditure

3.1.1 Simple linear regression

The model we employ is:

$$y_i|\mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$
$$\mu_i = \beta_0 + \beta_1 \text{cpi}_i + \epsilon \tag{2}$$
$$\beta_0 \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\beta_1 \sim \text{Normal}(0, 2.5) \tag{4}$$
$$\sigma \sim \text{Exponential}(1) \tag{5}$$

Where:

- y_i is the continuous outcome variable, representing the total expenditure in Canadian dollars in year i.
- cpi_i is the predictor variable, representing the annual Consumer Price Index (CPI) for year i

3.1.2 Multiple linear regression

In our second model, we added another predictor variable to employ a multiple linear regression model.

$$
\begin{aligned}
y_i|\mu_i, \sigma &\sim \text{Normal}(\mu_i, \sigma) && (6)\\
\mu_i &= \beta_0 + \beta_1 \text{cpi}_i + \beta_2 \text{income}_i + \epsilon && (7)\\
\beta_0 &\sim \text{Normal}(0, 2.5) && (8)\\
\beta_1 &\sim \text{Normal}(0, 2.5) && (9)\\
\beta_2 &\sim \text{Normal}(0, 2.5) && (10)\\
\sigma &\sim \text{Exponential}(1) && (11)
\end{aligned}
$$

Where:

- $income_i$ is the another predictor variable, representing the number of people earning an income in year i.

3.2 Model justification

From Figure 4 and Figure 5, we see a positive relationship in expenditure and CPI as the years pass. Implying that despite rising prices, household spending has continued to increase, specifically in the food and shelter categories. Due to this relationship, we anticipate the relationship is close to linear. We assume this relationship because food and shelter are necessities unlike recreational spending.

4 Results

Since linear models tend to focus on either inference or prediction (Wickham et al. 2019b), we have used it infer that expenditure will continue to increase as prices increase. Our model results are summarized in Table 1. The table compares our two models, each predicting the outcome variable (total expenditure), based on either CPI only, or CPI and the number of income earners.

In Table 1, we see the coefficient for the CPI is approximately 927, implying that for every unit increase in CPI, the total expenditure increases by approximately 927 units. In simple terms, CPI is a measure of inflation, and as inflation increases, expenditure increases. However, for the second model, we see a significantly smaller increase of 227 units in total expenditure. This illustrates that with the addition of the second predictor variable (number of income earners), we were able to get a more accurate representation because expenditure naturally increases

Table 1: Explanatory models of expenditure

	CPI	Add Income Earners
(Intercept)	−32 710.73	−59 621.55
	(9386.74)	(18 941.67)
cpi	927.08	227.39
	(73.60)	(429.83)
income		0.01
		(0.01)
Num.Obs.	10	10
R2	0.945	0.955
R2 Adj.	0.928	0.934
Log.Lik.	−88.432	−87.297
ELPD	−91.8	−91.2
ELPD s.e.	2.9	2.3
LOOIC	183.6	182.3
LOOIC s.e.	5.9	4.6
WAIC	182.4	180.8
RMSE	1503.15	1562.92

as the number of income earners increases, due to the assumption of lifestyle inflation, people will spend more money if they make more money. Table 1 also indicates the coefficient for income variable, which is quiet small, 0.01. However, this number did not worry us, as the scale is important in this context. The number of income earners is in the millions while the total expenditure is in the thousands. Since the coefficient is positive, we can predict that as the number of the income earners increase, the total expenditure would increase as well.

The first model provides simplicity and ease, while the second model includes another factor for review. The first model is the ideal fit as the standard errors are smaller. Policymakers can also make decisions solely on the impact of CPI alone. The number of income earners is relevant. However, from the coefficient for the income variable, we can conclude the average household income is not impacted too much. Expenditure is a result of many factors, household size, seasons, income range of the household, and the population and number of income earners definitely plays a role in how expenditure can increase.

5 Discussion

5.1 Findings

The data and variables were discussed in Section 2 and the Section 4 identified the results of our model in Section 3. Using all this, we identified our main findings. A common theory is that as inflation rises, people spend less. We conclude this theory is false because people have to spend more to survive, as found from our model results. We saw some decreases in recreation but overall, it increased. This could be do to the fact that technology and social media, has been widely used to make more money in less time, unlike a 9am to 5pm job. Thus, as prices increase, people must be making more money or borrowing money to make due an spend on necessities like food and shelter.

5.2 Bias

There are also other factors to consider that influence total expenditure. For example, unemployment rate and interest rates. For simplicity and due to data availability, we chose it to include two predictor variables only, since prices have a significant impact on spending. For the spending and CPI data, we used aggregate data to keep it consistent and broad. However, this can mask the variations in each dataset. For example, spending and CPI on entertainment and transportation look different than food and shelter. Thus, there can be misleading averages in a year. From the income data, we chose "all persons with income" which can undermine the experiences of marginalized vulnerable groups in the region.

5.3 Weaknesses

This paper focused explicitly on the total expenditure, inflationary changes over a period of several years, and the number of income earners in Ontario. Household spending data was only available from years 2010-2021. This is a relatively small interval so we had limitations for our model. However, it is still useful because a lot can change in a decade and if you look at data before 2010, it will not give you a clear visual of what is happening today. Thus, making it harder to make political or economical decisions that impact today's world.

5.4 Policy Implications and Future Research

It is important to find out whether people are earning more income or depleting their savings, or borrowing more than they can handle. With this information, policymakers can identify a clear problem, and that is the first step in developing a solution for the problem. If not, the region can face a significant economical collapse.

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.

Canada, Government of. n.d. *Statistics Canada*. https://www.statcan.gc.ca/en/start.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

von Bergmann, Jens, and Dmitry Shkolnik. 2021. *Cansim: Accessing Statistics Canada Data Table and Vectors*. https://CRAN.R-project.org/package=cansim.

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. 2019a. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

———, et al. 2019b. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Thomas Lin Pedersen, and Dana Seidel. 2023. *Scales: Scale Functions for Visualization*. https://CRAN.R-project.org/package=scales.