Analyzing Housing Affordability in the United States: The Role of Inflation, Interest Rates, and Supply*

Rising Inflation and Interest Rates Worsen Affordability, While Increasing Housing Supply Offers Relief

Krishna Kumar

November 26, 2024

This paper investigates housing affordability in the United States from 2006 to 2024 by constructing an Affordability Index that measures the number of work hours required to purchase a median-priced home outright. Using a multiple linear regression model, we analyze the relationships between housing affordability and three key macroeconomic variables: Consumer Price Index (CPI), housing supply, and interest rates. Our findings reveal that rising inflation and interest rates significantly reduce affordability, while increased housing supply mitigates this effect. These insights highlight the complex interplay of economic forces shaping the housing market and emphasize the importance of stabilizing inflation, increasing supply, and maintaining balanced monetary policy to improve housing accessibility for U.S. households.

1 Introduction

Housing affordability is a growing issue in the United States, with significant implications for economic stability and social well-being of individuals. Over the past two decades, escalating home prices, inflation, and shifting interest rates have exacerbated housing affordability issues, leaving many households unable to achieve home ownership. As policymakers seek solutions to these challenges, understanding the factors that drive housing affordability has become increasingly critical. This paper addresses this issue by focusing on the interplay of key macroeconomic variables—Consumer Price Index (CPI), housing supply, and interest rates—and their collective influence on housing affordability over time.

*Code and data are available at: https://github.com/RohanAlexander/starter_folder.

To capture the complexity of this issue, we develop an Affordability Index, a measure representing the number of work hours required to purchase a median-priced home outright. This index provides a clear, intuitive way to assess affordability trends int he long run and evaluate the impact of inflation, housing supply, and monetary policy on home ownership accessibility. Using multiple linear regression, we analyze how these factors interact and contribute to changes in affordability from 2006 to 2024. By systematically building and comparing three models, we examine the power of each predictor (marginally and collectively) and evaluate their significance in shaping affordability dynamics.

The findings of this paper reveal that inflation and interest rates significantly reduce housing affordability (i.e. increase the affordability index), while an increase in housing supply improves it. Rising CPI drives up the cost of living and housing prices, decreasing purchasing power. Similarly, higher interest rates inflate mortgage costs, further straining affordability. Conversely, increased housing supply alleviates upward pressure on prices, demonstrating its critical role in improving affordability. These results offer valuable insights for policymakers, highlighting the need for measures that stabilize inflation, increase housing stock, and balance monetary policy to enhance housing accessibility.

The remainder of this paper is organized as follows. Section Section 2 presents the data sources, variable descriptions, and exploratory analysis. Section Section 3 outlines the model details, setup, and justification. Section Section 4 highlights the key findings from the models and visualizations. Finally, Section Section 5 delves into the broader implications of these findings, addresses limitations, and suggests directions for future research. Moreover, Sections Section A and Section B of the appendix present the model diagnostics and a survey, sampling, and observation methodology that would be ideal for nature of data that we have used.

1.1 Estimand

The estimand in this paper is the Affordability Index, representing the hours of work required to purchase a median-priced home outright. This measure allows us to quantify and analyze the impact of macroeconomic variables on affordability in a clear and meaningful way, especially when the analysis focuses on the long run. By focusing on this estimand, we contribute to a deeper understanding of housing market dynamics and provide actionable insights for addressing affordability challenges.

2 Data

2.1 Overview

This report uses monthly economic data derived from the Federal Reserve Economic Data (FRED) database to examine the key factors influencing housing affordability in the United

States of America from March 2006 to September 2024. For this analysis, I have combined five data sets into one, which is then used for analysis. The data sets were downloaded using the `fredr` (Boysel and Vaughan 2021) package and loaded using the same library. The data were cleaned and analysed using the statistical programming language `R` (R Core Team 2023). Simulating, cleaning, testing, and graphing were done with the help of the following packages: `Tidyverse` (Wickham et al. 2019), `dplyr` (Wickham et al. 2023), `lubridate` (Grolemund and Wickham 2011), `testthat` (Wickham 2011), `readr` (Wickham, Hester, and Bryan 2024), `arrow` (Richardson et al. 2024), `ggplot2` (Wickham 2016), `here` (Müller 2020), `DiagrammeR` (Iannone and Roy 2024), '`DiagrammeRsvg` (Iannone 2016), `rsvg` (Ooms 2024), and `modelsummary` (Arel-Bundock 2022).

The Federal Research Economic Data provides informati on about the people and economy of the United States, with a goal to support economic growth, assist in making informed decisions, and enhance scientific knowledge. Housing affordability has emerged as a growing concern in the US, with people facing a shortage of houses and constantly increasing prices due to a variety of factors. These challenges, compounded by the COVID-19 pandemic, have left many families struggling to access affordable housing (Online 2024). Studying this data is essential in assessing the economic factors affecting housing affordability as well as the financial well-being of individuals. The data sets were all downloaded separately using the FRED API and then combined during the data cleaning.

2.2 Data Cleaning

The main aspect of the data cleaning was merging all five data sets into one data set. Moreover, the columns were renamed to remove any informal names and two new columns were created using the existing data sets: House Prices Growth (%) and Wage Growth (%). This was done using the simple formula of (House Price 2 - House Price 1)/House Price 1, calculating the percentage change. The same was done for average wages. This allowed us to analyse the growth rates year over year and get a sense of how fast each variable was growing through time. Additionally, another column was created titled 'Affordability Index.' This column was created by dividing the values of House Prices by the Average Wage in the same rows. This gave us an outcome of the number of hours an individual would have to work to buy a house outright, serving as a good indicator of housing affordability, especially when looking at the data through a long-term lens.

2.3 Data Limitations and Similar Data sets

While the data from FRED provides valuable insights into macroeconomic variables, it is not without limitations. Notably, the data on wages reflects average hourly earnings of all employees in the private sector, rather than total household income, which may overlook other forms of income, such as government transfers, investments, or multi-income households, that influence affordability. Additionally, the data is primarily aggregated at the national level,

limiting its ability to capture trends at the regional level in housing affordability, which can be substantial across states and metropolitan areas. Another consideration is that some variables are seasonally adjusted while others are not. However, over a long period of time, like the nearly two-decades used in this analysis, these seasonal adjustments are unlikely to impact the broader trends and implications this paper aims to study.

Similar data sets for all variables were readily available on the FRED database, along with other databases like the U.S. Census Bureau or Bureau of Labor Statistics. While those data sets might be able to provide regional or demographic data, the chosen variables were selected due to the monthly recording of data for a more detailed analysis. Despite these limitations, the FRED data sets turned out to be the best choice for an analysis fo this kind.

2.4 Variables of Interest

For analysis, only the required variables' data sets were downloaded from FRED. This paper decided to focus on five variables to assess the problem of unaffordable housing in the US: CPI, interest rate, housing supply, median housing prices, and average wages. They are described below:

- ***CPI*:** This variable represents the Consumer Price Index for All Urban Consumers, for all items in the U.S. City Average, which measures the average change over time in the prices paid by urban consumers for a market basket of consumer goods and services. Recorded monthly, this variable is measured in an index form, for which the base period is 1982-1984 where the value of the index sat at 100. Change in CPI from last year's CPI is one of the most common measures used to calculate inflation. This variable is seasonally adjusted and looks at ~ 88% of the population ("Consumer Price Index for All Urban Consumers: All Items in u.s. City Average_2024" 2024).
- ***Interest Rate:*** This variable refers to the Federal Funds Effective Rate, which is the interest rate at which depository institutions trade federal funds with each other overnight. This interest rate is the central interest rate in the US, influencing other interest rates such as the prime rate (what banks charge customers with higher credit ratings) and longer term rates (mortgages, loans, savings). This variable is recorded monthly, represented as a percentage, and is not seasonally adjusted ("Federal Funds Effective Rate_2024" 2024).
- ***Average Wage:*** This variable denotes the Average Hourly Earnings of All Employees in the Total Private Sector. It includes premium pay for overtime and late-shift work. This was chosen over wage rates as this measures the actual return to a worker for a set period of time. It excludes benefits, irregular bonuses, retroactive pay, and payroll taxes paid by employers. This variable is measured in US Dollars per hour, recorded monthly, and is seasonally adjusted ("Average Hourly Earnings of All Employees, Total Private_2024" 2024).

- ***House Price:*** This variable captures the Media Sales Price for New houses Sold int he United States. The media was chosen over the average prices to remove the effects of any outliers. This variable is recorded monthly, measured in US dollars, and is not seasonally adjusted ("Median Sales Price for New Houses Sold in the United States_2024" 2024).
- ***House Supply:*** This variables measures the Monthly Supply of New Houses in the United States. Measured in 'month's supply,' this refers to the ratio of new houses for sale to new houses sold. The month's supply indicates the size of the new for-sale inventory compared to the number of new houses currently being sold. This variable is recorded monthly and is seasonally adjusted ("Monthly Supply of New Houses in the United States_2024" 2024).
- ***House Price Growth:*** This variable represents the year-over-year percentage growth in the Median Sales Price for New Houses Sold in the United States. It is calculated by taking the difference between the median house price at the end of the year and the beginning of the year, dividing this difference by the value at the beginning of the year, and multiplying by 100 to express the result as a percentage. This highlights the annual rate of change in house prices, helping to track trends in housing market inflation or deflation.
- ***Wage Rate Growth:*** This variable reflects the year-over-year percentage growth in the Average Hourly Earnings of All Employees in the Total Private Sector. The calculation follows the same approach as house price growth: the difference between wages at the end of the year and the beginning of the year is divided by the wage value at the start of the year, then multiplied by 100 to express the result as a percentage. This variable captures how wages change annually, offering insights into trends in earning power.
- ***Affordability Index:*** This variable measures the ratio of the Median Sales Price for New Houses Sold in the United States to the Average Hourly Earnings of All Employees in the Total Private Sector. It represents how many hours of work are required to afford a median-priced home. A higher value indicates reduced affordability, as workers would need to work more hours to buy a home.

2.5 Measurement

- ***CPI*:** The CPI is based on prices for food, clothing, shelter, and fuels; transportation fares; service fees (e.g., water and sewer service); and sales taxes. Prices are collected monthly from about 4,000 housing units and approximately 26,000 retail establishments across 87 urban areas. To calculate the index, price changes are averaged with weights representing their importance in the spending of the particular group. The index measures price changes (as a percent change) from a predetermined reference date.

- ***Interest Rate:*** The Federal Open Market Committee (FOMC) meets eight times a year to determine the federal funds target rate. These rates are then announced to the public.
- ***Average Wage:*** This is measured through the Current Employment Statistics (CES) program, which surveys approximately 119,000 businesses and government agencies monthly. Employers report total gross payroll and hours worked for all private-sector employees during the pay period including the 12th of the month. Average Hourly Earnings are calculated by dividing total payroll by total hours worked, excluding irregular bonuses or benefits, and aggregated using the NAICS classification system to provide an accurate measure of wage trends across industries.
- ***House Price***: This is measured through data collected by the U.S. Census Bureau and the Department of Housing and Urban Development (HUD). This data is based on surveys of new single-family houses sold, with the price representing the middle value when all houses sold during a given period are ordered by price. The calculation excludes multifamily units and focuses on single-family homes, ensuring the measure reflects trends in the mid-range of housing markets without being skewed by extreme high or low prices.
- ***House Supply***: This is measured as the ratio of the number of new houses available for sale to the number of new houses sold, expressed in months. Data is collected through the New Residential Sales Survey, conducted jointly by the U.S. Census Bureau and HUD, which tracks both housing inventory and sales rates. This measure reflects how long the current inventory of new houses would last at the current sales pace, assuming no new inventory is added. The measure is calculated by dividing the total inventory of new homes available for sale by the current monthly sales rate. For example, if there are 400,000 houses for sale and 40,000 houses are sold in a month, the monthly supply is \$400,000/40,000 = 10 months. A higher value suggests excess inventory (a buyer's market), while a lower value indicates limited inventory (a seller's market).

2.5.1 Consumer Price Index Trends Over Time

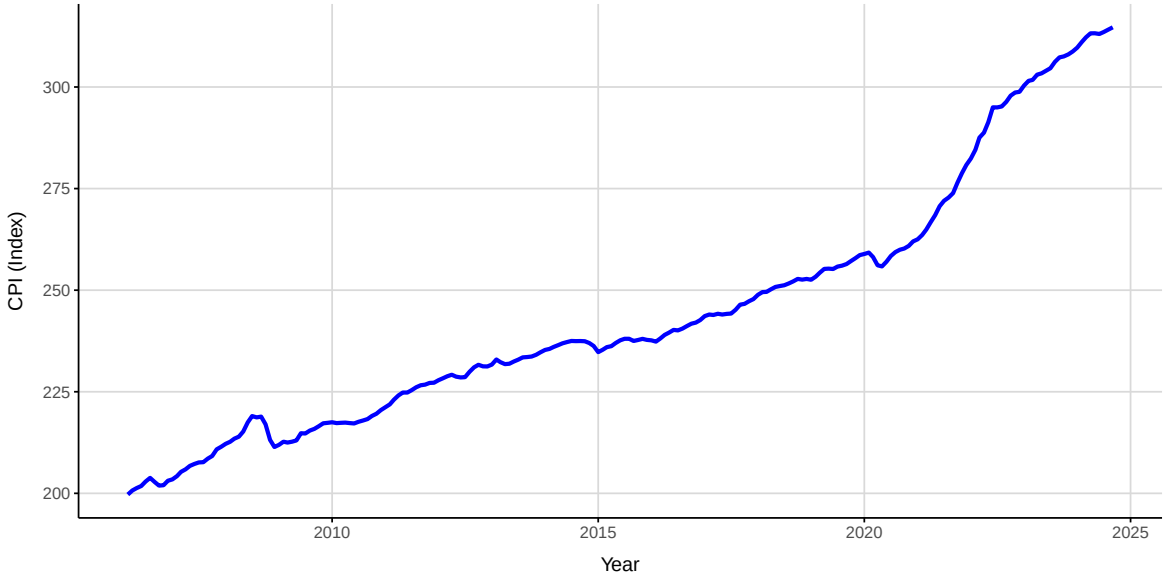

Figure 1: Trend in Consumer Price Index (CPI) from 2008 to 2024, highlighting a steady rise in inflation with accelerated growth post-2020.

Figure Figure 1 shows a clear pattern of a gradual increase from 2006 to 2024, with a clear spike in 2007 due to the financial crisis and a drop immediately after 2020 as a result of the COVID-19 pandemic. We can also see a notable acceleration in the post pandemic period CPI where the trend line becomes much steeper. This highlights periods of economic expansion, but also signals rising costs for consumers.

Inflation, as measured by CPI, directly impacts housing affordability. As inflation rises, the cost of living increases, including housing prices. Higher inflation often prompts the Federal Reserve to raise interest rates (which we will see in later graphs), making mortgages more expensive and reducing affordability for potential home buyers. Furthermore, when inflation erodes purchasing power, individuals have less disposable income to allocate toward home ownership, exacerbating affordability challenges. Historical relationships between CPI and housing prices reveal that rising inflation can contribute to increased home prices due to higher costs for construction materials and labor (Nguyen 2023).

2.5.2 Interest Rates Trends Over Time

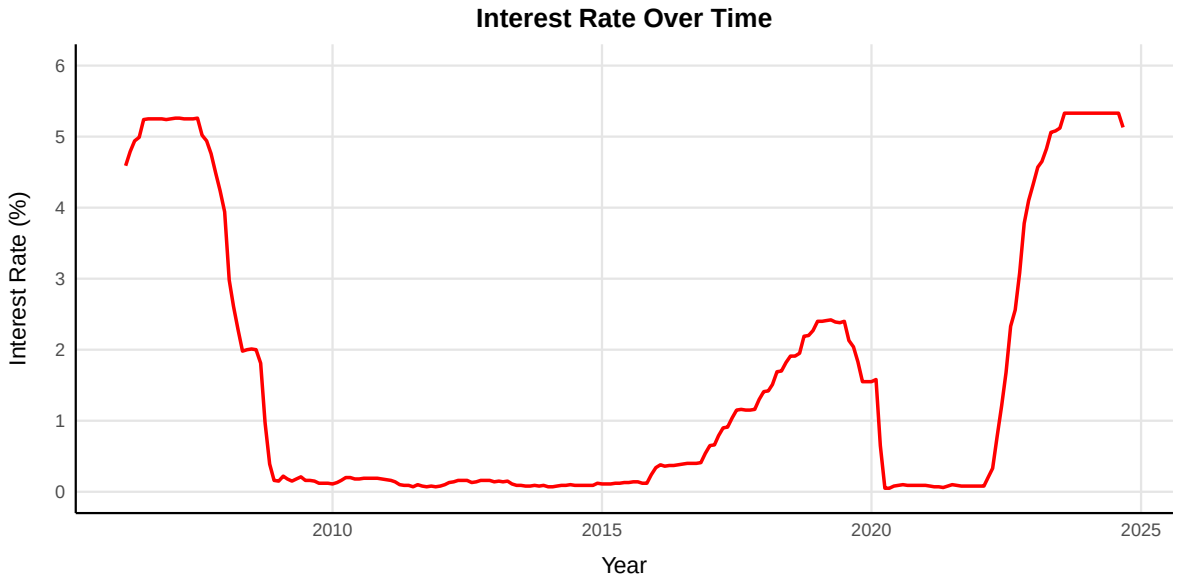

Figure 2: Historical trend of interest rates from 2008 to 2024, showcasing significant declines post-2008 financial crisis and sharp increases in recent years to combat inflation.

Figure Figure 2 shows the irregular trends of the interest rate from 2006 to 2024. The interest rates saw a drastic drop during the financial crisis which began in 2007. From 2008 to 2015, interest rates remained historically low, hovering near zero due to monetary policy responses to the Great Recession. A gradual increase began in 2016, reflecting economic recovery and attempts to normalize monetary policy, before another sharp decline in 2019 as the COVID-19 pandemic spurred aggressive monetary easing to support the economy. Post-2022, we can see a dramatic rise in interest rates, corresponding to central banks' efforts to combat inflationary pressures. This sharp increase represents one of the fastest monetary tightening cycles in recent history.

Regarding housing affordability, these trends are crucial. Low interest rates, such as those seen during 2008–2015 and 2020–2021, generally reduce borrowing costs, making mortgages more affordable and stimulating demand in the housing market. Ironically, as we will see in Figure 3, during these times when borrowing becomes cheaper, we have seen housing supply fall drastically. Moreover, the sharp increases in interest rates post-2022 significantly raise mortgage rates, reducing affordability for new buyers and potentially cooling housing market activity, and of course, in these times, we have seen an increase in housing supply. These trends underline the strong interplay between monetary policy and housing market dynamics, highlighting the importance of interest rate policies in influencing housing affordability over time

2.5.3 Housing Supply Over Time

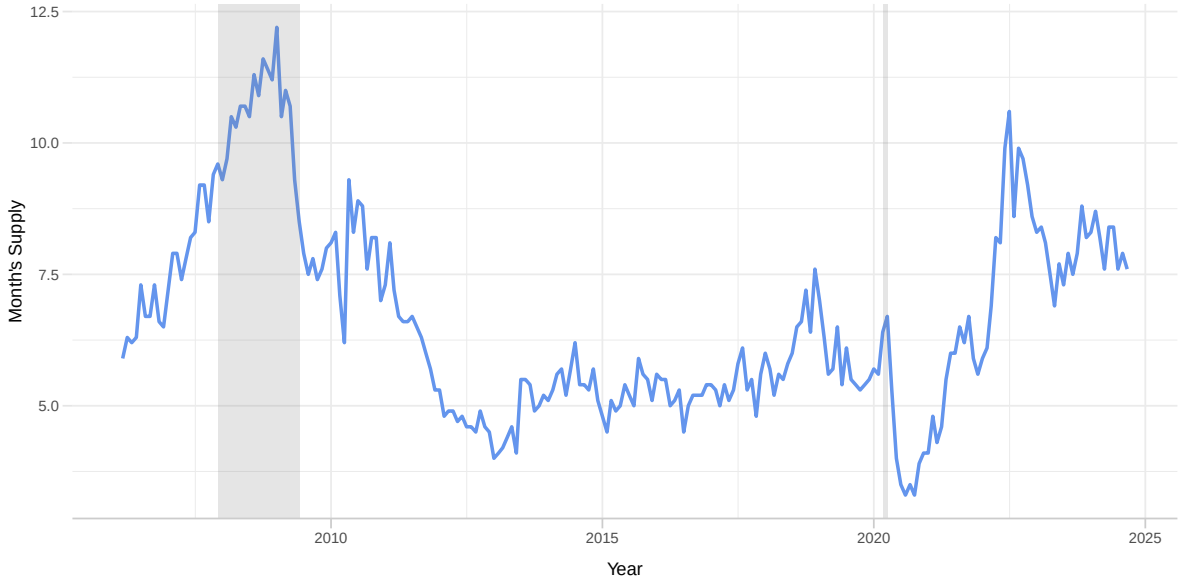

Figure 3: Monthly housing supply trends from 2008 to 2024, highlighting peaks during the 2008 financial crisis and the COVID-19 pandemic, followed by declines as markets adjusted.

Figure Figure 3 highlights sharp fluctuations in housing supply in the past 19 years, with recessions marked by the shaded areas. In the aftermath of the 2007–2008 financial crisis, housing supply spiked to over 12 months, reflecting a surplus of available homes due to decreased demand and widespread foreclosures. However, as the economy began recovering, supply steadily declined, hitting a low point around 2013. This downward trend in supply coincided with robust demand, driven by low interest rates and favorable borrowing conditions.

During the COVID-19 pandemic, a combination of supply chain disruptions, labor shortages, and heightened demand for housing led to a sharp decline in supply, reaching historic lows. Post-2020, the supply began to rise again, likely due to reduced affordability as a result of rising interest rates (as seen in Figure Figure 2) and cooling housing market activity. However, it reached a high point in the first half of 2022 and then kept decreasing drastically till 2024.

2.5.4 Wage Growth vs House Prices

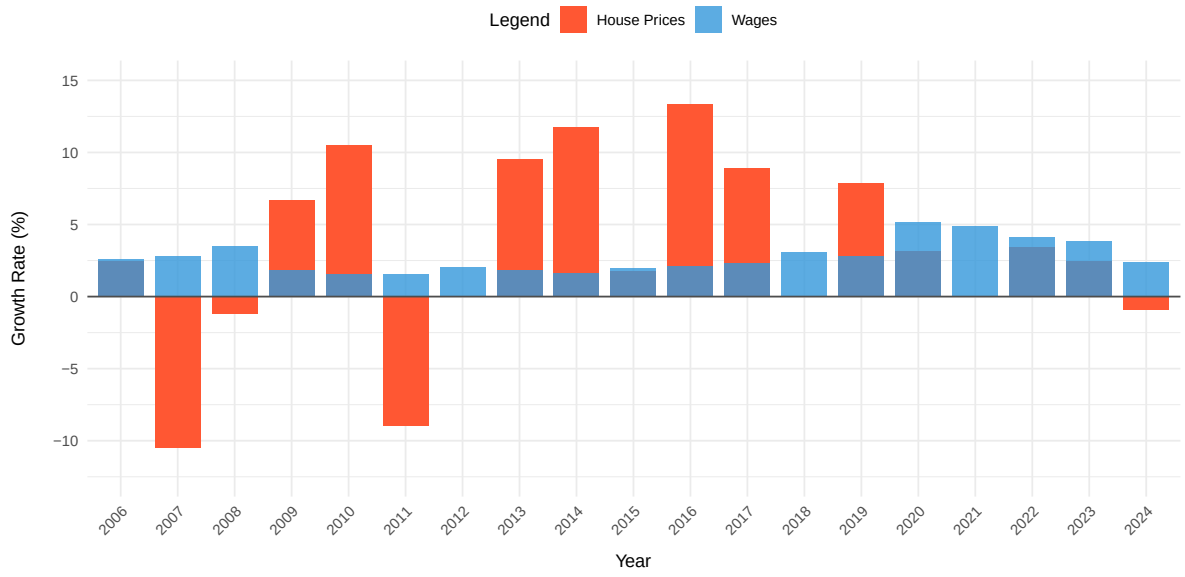

Figure 4: Comparative growth rates of house prices and wages from 2006 to 2024, showing periods where house price increases significantly outpaced wage growth, particularly during housing market recoveries and economic expansions.

Figure Figure 4 compares the annual growth rates of house prices and wages from 2006 to 2024. The data highlights a persistent gap between the two metrics, with house prices outpacing wage growth in 5 of the 19 years analysed. During the financial crisis of 2007–2008, house prices saw a sharp drop, with growth rates falling below -10%, while wage growth remained relatively stable. Post-2011, house prices recovered, showing strong growth from 2012 to 2016, significantly exceeding wage growth during this period. This divergence reflects a period of increasing housing demand, driven by low interest rates and insufficient housing supply, as seen in Figure Figure 3.

Notably, wage growth has consistently lagged behind house price growth, contributing to worsening housing affordability over the years. For example, during the COVID-19 pandemic in 2019, while wage growth remained positive, house prices surged, reflecting the impact of supply shortages and pandemic-induced demand. Post-2022, a notable shift occurred with house price growthdeclining due to rising interest rates and reduced affordability, while wage growth maintained a steady pace. These trends emphasize the widening affordability gap and the complex interplay between wages, housing prices, and macroeconomic conditions, highlighting the challenges faced by households in balancing income growth with escalating housing costs.

2.5.5 Affordability Index

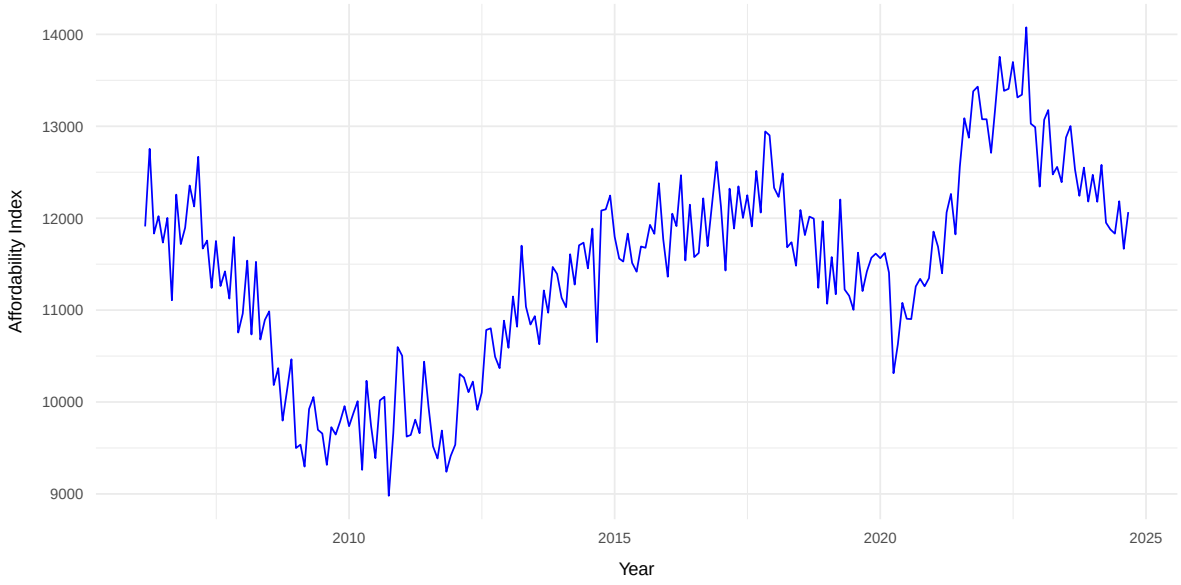

Figure 5: Trends in the Housing Affordability Index from 2006 to 2024, highlighting periods of increased affordability during economic downturns and decreased affordability during housing market peaks.

Figure Figure 5 presents the trends in the affordability index from 2006 to 2024, illustrating fluctuations in the cost of housing relative to wages. The affordability index is a key metric that represents the number of work hours required to afford a house, with higher values indicating greater financial strain on individuals. The index shows notable volatility over time, particularly during key economic events. From 2008 to 2011, the index declined significantly, reflecting improved affordability during the global financial crisis, likely due to falling house prices and stabilizing wages. However, affordability worsened from 2012 onwards, as the index steadily increased, peaking around 2020. This peak coincides with rising house prices during the pandemic-driven housing boom, requiring more work hours to afford a home. Post-2020, the index shows a downward trend, reflecting corrections in housing prices amid rising interest rates.

3 Model

In this section, we develop and evaluate a series of models to understand the factors influencing housing affordability. The major challenge lies in selecting a model which is balanced in complexity and provides insightful results. We will be analyzing how CPI, housing supply, and interest rates affect the housing affordability in three separate models, adding one after

the other. The goal of out modelling strategy is to find the effect of each of the predictor variables on the affordability index and their strengths.

Because housing prices or wage rates alone can't determine affordability, we created an affordability index to better measure affordability. This index has a simply formula where we divide housing prices by average wages, leading to a number that represents how many hours an individual would need to work to buy a house outright.

We begin with a single predictor, CPI, to assess it's impact on housing affordability. Then, we add housing supply for model 2 and interest rates for model 3, arriving at a full model that shows us the effect of each of these variables marginally, as well as collectively. This step wise approach helps ensure that the model remains interpretable. Moreover, performance measures such as R2, Adjusted R^2, AIC, BIC, and RMSE were used to evaluate model performance.

Here we briefly describe the multiple linear regression model used to investigate… Background details and diagnostics are included in Section A of the Appendix.

3.1 Model set-up

We aim to model the Affordability Index, which represents the number of hours of work required to purchase a house outright, as a function of key economic variables such as interest rates and inflation (measured by CPI).

$$y_i = y_i = \beta_0 + \beta_1 \cdot \text{CPI}_i + \beta_2 \cdot \text{HouseSupply}_i + \beta_3 \cdot \text{InterestRate}_i + \epsilon_i$$

Where:

- y_i is the affordability index for observation iii, which measures the ratio of house prices to average wages, indicating housing affordability.
- β_0 is the intercept, representing the expected value of the affordability index when all predictors are zero.
- β_1 captures the effect of the CPI (Consumer Price Index) on the affordability index, reflecting inflation.
- β_2 captures the effect of the Housing Supply on the affordability index.
- β_3 captures the effect of the Interest Rate on the affordability index
- ϵ_i represents the error term, assumed to follow a normal distribution with mean 0, accounting for unexplained variability in the affordability index.

3.2 Model Justification

The linear regression model was chosen for this analysis to understand how each predictor variable affects affordability of housing. The affordability index is key to understanding patterns of affordability in the housing market; This is something average wages or house prices cannot determine along. CPI, housing supply, and interest rates all have impacts on housing affordability we expect to see a strong positive correlation between the affordability index and CPI and interest rates. Meaning that as these variables increase, the affordability index increases as well, indicating worsening affordability. However, we expect to see a negative relationship between the affordability index and housing supply. We also aim to measure which of these variables has a stronger impact on housing affordability.

An increase in interest rates increases the mortgage rates in turn and decreases housing affordability. An increase in CPI represents an increase in inflation, which then usually leads to an increase in housing prices. Lastly, a decrease in the housing supply leads to a shortage of homes and therefore, leads to increased housing prices and lower affordability. These factors work alone and together, to influence the financial burden on individuals. We decided not to choose wage growth or house price growth as predictor variables in our analysis as those would have a direct conflict with the outcome variable, and would lead to an inaccurate model.

We opted for a linear regression model due to its capacity to quantify the marginal, and collective, effects of each predictor (CPI, housing supply, and interest rates) on housing affordability. This structure is well suited to match the topic and context of our analysis.

All modelling was conducted using the Base R package, specifically using the lm() function from the statistical package for linear regression analysis.

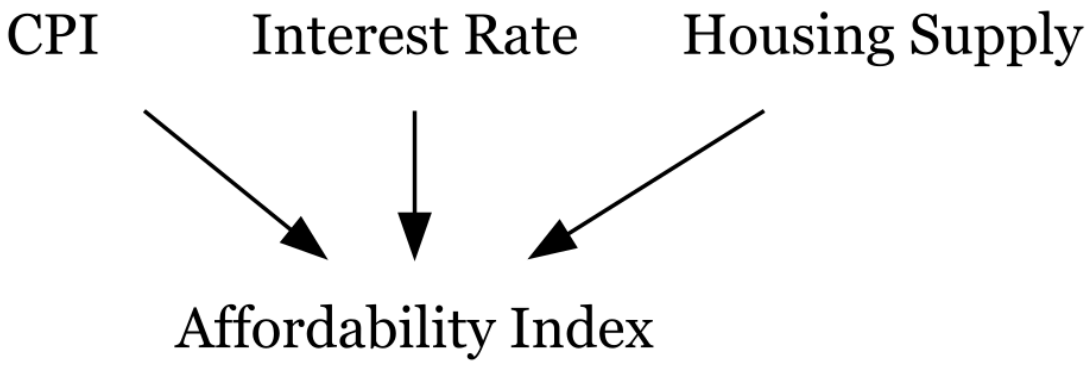

Figure 6: Factors influencing the Affordability Index shown as direct contributors

Table 1: Model Performance Summary showing improved fit and accuracy as Housing Supply and Interest Rates are added, with R^2 increasing from 0.370 in Model 1 to 0.471 in Model 3, and RMSE decreasing by over 70 points

	Model 1	Model 2	Model 3
(Intercept)	6113.14	6411.21	7499.80
	(469.64)	(518.18)	(507.91)
CPI (Index)	21.75	21.65	18.24
	(1.91)	(1.91)	(1.84)
House Supply (Month's Supply)		−41.41	−125.59
		(30.68)	(31.27)
Interest Rate (%)			198.23
			(31.41)
Num.Obs.	223	223	223
R2	0.370	0.375	0.471
R2 Adj.	0.367	0.370	0.464
AIC	3649.2	3649.4	3614.1
BIC	3659.4	3663.0	3631.2
Log.Lik.	−1821.611	−1820.691	−1802.059
F	129.840	66.073	65.101
RMSE	853.82	850.31	782.15

4 Results

Our results are summarized in Table 1.

Table Table 1 summarizes the performance metrics for our three models, with progressively added variables. Model 1, which only uses CPI as a predictor, achieves an R^2 of 0.370, indicating that CPI alone explains about 37% of the variance in the affordability index. Model 2 uses housing supply as an added predictor and achieves only a 0.005% increase in R^2 , indicating that housing supply has a very minute impact on the variation of the affordability index. Moreover, model 2 only sees a very small reduction in the RMSE (850.31) showing a status quo in predictive accuracy and model fit. This allows us to conclude that housing supply, when combined with CPI does not produce an impact that is statistically significant.

Model 3, which incorporates all three predictor variables, shows the largest improvement, with R^2 rising to 46.4% and RMSE dropping to 782.15, reflecting a stronger model fit. Moreover, model 3 also sees the lowest values of AIC and BIC that indicate that model 3 is the best

model as it balances fit and complexity most effectively. This progression highlights the benefit of adding contextual variables like housing supply and interest rates to better capture the accuracy of the model, as well as analyzing the marginal impacts of each predictor variable.

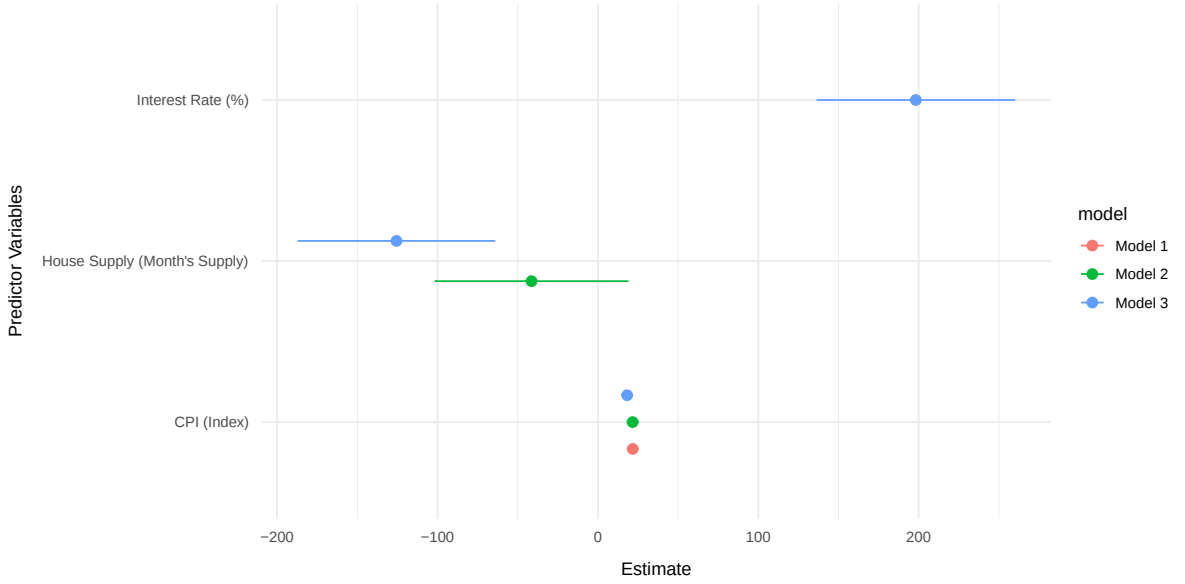

Figure 7: Coefficient estimates with 95% confidence intervals across three models, showing that CPI and interest rate positively affect the affordability index, reducing affordability, while an increase in housing supply improves affordability.

Moreover, Figure 7 visually represents the coefficients of the predictors in the three models, with their corresponding confidence intervals. Each point represents the estimated effect size (coefficient) of a predictor, and the lines extending from each point indicate the range of uncertainty (confidence intervals).

The CPI coefficients remain positive for all three models, indicating that higher inflation (measured by CPI) is associated with a higher affordability index (i.e. less affordable housing). However, we see that the coefficients for housing supply are in the negatives for both models it's used as a predictor in. Suggesting that an increase in housing supply improves affordability, which is consistent with our expectations. Lastly, the coefficient for interest rate is strongly positive, indicating that higher interest rates significantly increase the affordability index (worsening affordability), likely due to higher mortgage payments. The wider confidence intervals of housing supply and interest rate depict uncertainty, while the absence of a confidence interval in CPI represents more statistically precise estiamtes. The results we have seen from Table 1 and Figure 7 are completely consistent with our broad economic expectations, however, this model has room for improvement in its predictive accuracy, model fit, and certainty.

5 Discussion

5.1 Overview: What Is Done In This Paper?

This paper investigates housing affordability in the United States over the period of 2006–2024 by developing a Affordability Index. The index measures the number of hours of work required to purchase a median-priced home and serves as a metric for gauging the accessibility of home ownership across time, especially when analysed through a long-term lens. Using multiple linear regression, this paper examines the influence of key economic variables—Consumer Price Index (CPI), housing supply, and interest rates—on housing affordability. These factors were chosen because they capture inflationary pressures, supply constraints, and monetary policy, which are all fundamental drivers of housing market dynamics.

To conduct the analysis, five data sets from FRED were downloaded, compiled, and processed to align temporal and unit measurements. The model specifications were carefully designed to progressively evaluate the contribution of each predictor, starting with CPI as a standalone variable, followed by the inclusion of housing supply, and finally incorporating interest rates. This approach allowed for a clearer understanding of how these variables individually and collectively influence affordability. The results offer critical insights into how shifts in macroeconomic factors affect home-buying accessibility and shed light on the broader housing market trends over the last two decades.

5.2 Key Findings

5.2.1 Relationship Between CPI (Inflation) and the Affordability Index

The positive and statistically significant relationship between CPI and the affordability index confirms that higher inflation leads to reduced affordability. Inflation, as measured by the CPI, directly correlates with rising housing costs because construction materials, labor, and overall living expenses increase. As CPI rises, potential home buyers face higher costs not only for housing but also for other goods and services, reducing their disposable income for mortgage payments. This finding highlights the economic strain placed on individuals during periods of inflation and highlights the need for policies that mitigate inflationary pressures in housing markets.

5.2.2 Relationship Between Housing Supply and the Affordability Index

The negative coefficient associated with housing supply indicates that an increase in available housing improves affordability. When housing supply grows, competition among buyers decreases, stabilizing or reducing home prices. This relationship aligns with economic theory,

which states that greater supply relative to demand leads to price reductions. However, the relatively modest effect size, from our model, suggests that while increasing supply is a necessary condition for improving affordability, it may not be sufficient on its own without addressing broader structural issues, such as zoning restrictions and construction costs. However, this could also be due to the limitations of the model.

5.2.3 Relationship Between Interest Rates and the Affordability Index

Interest rates exert the most substantial impact on affordability among the predictors. The strong positive coefficient indicates that rising interest rates significantly worsen affordability by increasing mortgage costs, as interest rates affect other rates like the mortgage rates as well. For instance, a higher interest rate raises monthly payments, reducing the number of households that qualify for home loans. Moreover, looking at this from the perspective of the potential hoome buyer, it limits their choices and they are often forced to decide against buying a home and renting. This finding emphasizes the critical role of monetary policy in shaping housing market accessibility. Policymakers must carefully consider the downstream effects of interest rate hikes, especially in a high-inflation environment where affordability is already strained. While interest rates are often adjusted to combat the problem of growing inflation in an economy, its subsequent effects that affect the quality of life of citizens need to be considered as well.

5.3 Limitations

5.3.1 Data Aggregation

The data sets were aggregated at the national level, which limits the analysis' ability to capture regional variations in housing affordability. Housing markets in urban centers like San Francisco or New York operate under vastly different conditions than rural markets, and these distinctions are obscured in a national-level analysis. Studying the differences in housing affordability in rural vs urban areas as well as east vs west coast would be a very interesting, relevant, and topical concept to understand the true nature of housing affordability in the US. Future studies should aim to disaggregate data to better account for these geographic differences and further the analysis.

5.3.2 Model Fit and Scope

The adjusted R^2 values for the models range between 37% and 46%, indicating that the predictors explain less than half of the variation in the Affordability Index. This suggests that other unobserved factors—such as demographic shifts, wage inequality, government policies and support, regional income distributions, and renting and the supply of apartment complexes—play significant roles in determining housing affordability. For instance, regional wage disparities

or the availability of housing subsidies could significantly influence affordability but were not included in this analysis.

5.3.3 Exclusion of Other Influential Variables

As the adjusted R^2 values for the model were relatively low, we can decipher that there are several other factors that account for the remaining variance in the affordability index. Key factors such as household savings rates, employment stability, government intervention programs (e.g., tax credits or housing subsidies), rental properties, apartment complexes, condos, planning restrictions, and regional variations were not included due to data limitations. These factors likely play a critical role in shaping housing affordability but remain unexplored in this analysis. This, in no way, is a negative point of this paper, rather, it serves as a suggestion for extension of this analysis in further studies where we could understand the problem of housing affordability even more accurately.

5.4 Future Steps

5.4.1 Incorporating Additional Predictors

Expanding the model to include variables such as wage growth, population density, regional housing prices and income, and construction costs could enhance its explanatory power. For instance, wage growth relative to inflation could provide deeper insights into purchasing power, while construction costs could shed light on the supply-side challenges of the housing market. Conducting state- or city-level analyses would allow for a more granular understanding of housing affordability trends. This approach could highlight regional disparities and inform localized policy interventions. For example, comparing affordability in high-cost metropolitan areas to that in rural regions could reveal targeted strategies for improving housing access and government intervention.

5.4.2 Exploring Non-Linear Relationships

Future studies could employ non-linear modeling techniques, such as machine learning or interaction models like a Bayesian model, to capture more complex dynamics between variables. For example, the interaction between interest rates and inflation may have compounding effects on affordability that are not captured by a linear framework.

5.4.3 Evaluating Policy Interventions

Research focusing on the effectiveness of government programs, such as housing subsidies, rent control, or first-time home buyer incentives, could provide actionable recommendations for policymakers. Evaluating these interventions in the context of affordability trends would help determine which policies have the most significant impact.

In conclusion, this paper contributes to the broader understanding of housing affordability by analyzing the interplay of inflation, housing supply, and interest rates. The findings emphasize the critical role of macroeconomic variables in shaping housing accessibility and highlight areas where policymakers can intervene to mitigate affordability challenges. While the analysis has its limitations, it lays the groundwork for future research to build more comprehensive models and address the pressing issue of housing affordability in the United States.

Appendix

A Model Diagnostics

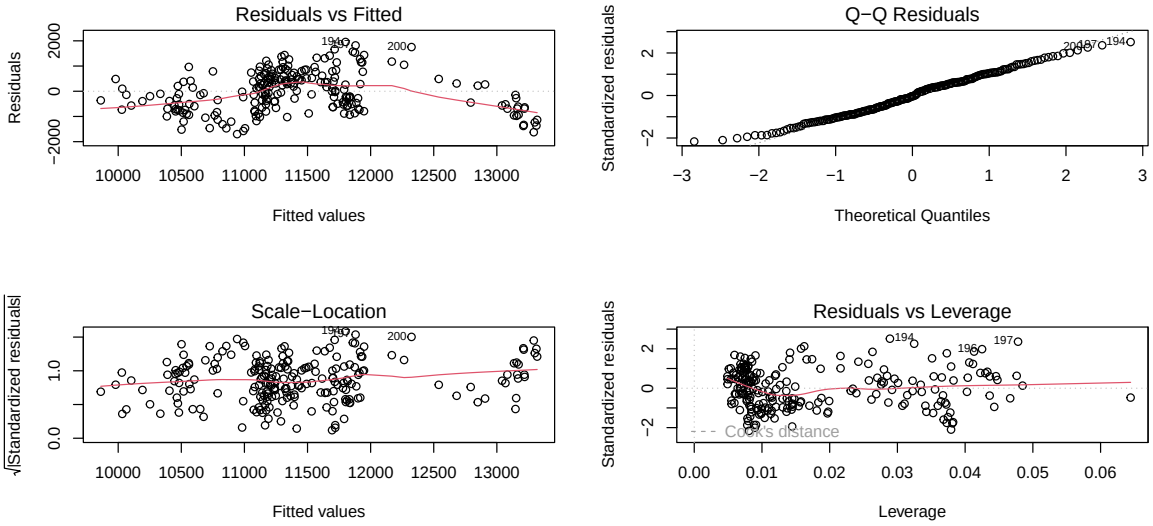

Figure 8: Diagnostic Plots of Model 3 show that assumptions are largely met, with slight deviations

The diagnostic plot for Model 3 is given in Figure 8 and we notice the following:

1. **Residuals vs. Fitted**: The residuals display a somewhat random scatter around the zero line, indicating that the linearity assumption is largely met. However, slight deviations suggest potential minor non linearity, particularly in specific ranges of fitted values. This indicates room for improvement but does not severely undermine the model's reliability.
2. **Q-Q Plot**: The residuals follow a relatively straight line, indicating that the assumption of normality is generally satisfied. Small deviations at the tails are visible, but these are minor and unlikely to significantly impact inference or predictions.
3. **Scale-Location Plot**: The residuals appear to be evenly spread, demonstrating that the assumption of homoscedasticity (constant variance of errors) is broadly met. There are slight variations in spread, but these are relatively minor and unlikely to skew results.
4. **Residuals vs. Leverage**: A few points with slightly higher leverage are observed, but their Cook's distance values indicate that they are not excessively influential. These points warrant attention but do not dominate the overall model.

B Idealized Methodology

B.1 Limitations with FRED Methodology

The Federal Reserve Economic Data's (FRED) data sets, while comprehensive in providing macroeconomic indicators, have several limitations when assessing housing affordability. First, it lacks granularity at the household level, failing to capture individual variations in financial burdens or regional disparities. Second, FRED focuses on objective metrics like house prices, wages, and economic variables, but omits subjective indicators such as perceived affordability or financial stress felt by individuals. Third, the data is often aggregated, potentially obscuring critical sub populations like low-income households or first-time home buyers. These gaps emphasize the need for complementary data collection methods that can provide a more holistic understanding of housing affordability.

B.2 Overview

To address the gaps identified in the FRED data sets and collection methods, we propose a comprehensive survey methodology aimed at capturing a clearer view of housing affordability in the United States. This approach integrates best practices from existing housing surveys, such as the Canada Mortgage and Housing Corporation's (CMHC) Social and Affordable Housing Survey ("Social and Affordable Housing Survey — Rental Structures Methodology_2023" 2023), and emphasizes both objective metrics and subjective experiences. The methodology focuses on diverse sampling techniques, inclusive recruitment strategies, and solid data validation processes to ensure the collection of reliable and representative data.

B.3 Sampling Method

To ensure that the survey captures a wide range of experiences and perspectives regarding housing affordability, we propose a **stratified multi-stage sampling design**. This design breaks down the population into smaller, more manageable groups and ensures that each group is properly represented in the final sample. Here's how it works:

1. **Stratification**: The first step divides the population into groups, or "*strata*," based on key characteristics such as geographic region, income level, type of housing (renter or homeowner), and urban or rural residence. For example, one stratum might consist of low-income renters in urban areas, while another might consist of middle-income homeowners in rural areas. By dividing the population in this way, we ensure that all important subgroups are included in the survey.

2. **Primary Sampling Units (PSUs)**: Within each stratum, large areas such as counties or cities are chosen as primary sampling units. These are selected using a method called "probability proportional to size" (PPS), which means areas with larger populations are more likely to be included. For example, If Los Angeles has a population of 4 million and a small rural area has a population of 40,000, the probability of selecting Los Angeles as a PSU is 100 times greater than that of the rural county. This is done in order to keep the probabilities of selection equal to the population sizes to give each area a proportinoally equal opportunity and representation.

3. **Secondary Sampling Units (SSUs)**: After the PSUs are chosen, smaller areas within them—such as neighborhoods—are selected. Again, PPS sampling is used to ensure that larger neighborhoods have a better chance of being included, due to the higher sampling frame. For example, within Los Angeles, densely populated neighborhoods like Downtown LA might have a higher probability of being chosen than sparsely populated suburban neighborhoods. However, suburban neighborhoods still have a chance to be included to ensure a balanced representation.

4. **Addressing Rural/Small Town Representation**:
   While PPS is efficient, it risks under representing rural areas as they often have lower populations and hence, a lower chance of being selected. These areas often face distinct housing affordability challenges. To mitigate this:

   - **Separate Strata for Rural Areas:** A separate strata will be constructed for rural areas and areas with smaller populations, from which PSUs and SSUs can be selected. We did mention this in the stratification section, but we must emphasize that we aim for equal representation of all areas and demographic types.
   - **Oversampling Rural Areas**: Rural PSUs can be over sampled within their strata, ensuring sufficient representation despite their smaller population sizes.
   - **Weighted Analysis**: Post-survey, weights can be applied to adjust for the lower selection probability of rural areas. For example, responses from a small rural neighborhood might be weighted more heavily than those from a densely populated city.
   - **Quota Sampling**: Quotas can be set to guarantee that a minimum percentage of the sample comes from rural PSUs, ensuring balance between urban and rural perspectives.

5. **Household Selection**: Finally, individual households are randomly selected from the chosen neighborhoods. This ensures that the survey captures a variety of household experiences within each subgroup.

This process creates a sample that represents the diversity of the U.S. population while also allowing for deeper insights into specific groups. For example, by oversampling low-income households or areas with high housing costs, the survey can provide detailed data about the

populations most affected by housing affordability challenges. This method is inspired by similar approaches used in large-scale housing surveys, such as the CMHC's Social and Affordable Housing Survey, which ensures a representative sample across diverse regions and demographics ("Social and Affordable Housing Survey — Rental Structures Methodology_2023" 2023).

B.4 Recruitment Strategy

A well-designed recruitment strategy is crucial for ensuring a diverse and representative sample for the survey. To achieve this, a **mixed-mode approach** will be used, employing multiple methods to contact participants and collect responses. This strategy not only maximizes response rates but also ensures that no group is excluded due to accessibility or technological barriers. To further maximize response rates and participation, we will hope to include a small compensation (small cash amount, gift cards, etc) for completing the survey, given that we can gather sufficient funding. Here's how the process works:

- **Mail Invitations**:
  The first step in reaching participants is through mailed invitations. These physical letters will provide an introduction to the survey, explaining its purpose, importance, and instructions for participation. Including a personal touch like mail ensures that even households without internet access or those less tech-savvy are aware of the survey. For instance, older adults who may not regularly use the internet are more likely to respond to mailed materials.

- **Online Surveys**:
  Online surveys will serve as the primary method for data collection because they are cost-effective, quick, and accessible to most people. Participants can simply follow a link from the mailed invitation or QR code to access the survey on their computers or smartphones. Online surveys also make it easier to process and analyze data efficiently, reducing both time and human error. This method is particularly effective for tech-savvy demographics, such as younger adults. The platform for these surveys will be Google Forms.

- **Telephone Interviews**:
  To accommodate households that lack internet access or prefer not to participate online, telephone interviews will be conducted. Trained interviewers will call participants and guide them through the survey, ensuring that they can share their responses comfortably. This approach is particularly helpful for older adults, rural households, and low-income groups who may face barriers to online participation. It ensures that no one is left out due to lack of internet or unfamiliarity with digital tools. All necessary steps will be taken to identify the household and individuals correctly to ensure full accuracy in data collection.

- **In-Person Interviews**:
  In areas with low response rates or where other methods prove ineffective, in-person

interviews will be conducted. Trained interviewers will visit selected households to collect responses face-to-face. This method is particularly useful in remote areas or among populations that might distrust online or phone methods. For example, face-to-face interactions often lead to higher response rates among communities where personal rapport is valued.

This mixed-mode approach is inspired by the Canadian Mortgage and Housing Corporation's (CMHC) survey methodology, which emphasizes using diverse data collection techniques to enhance data quality and fully represent all types of individuals and households. It ensures that every demographic—urban or rural, young or old, digitally connected or not—is represented in the survey results.

B.5 Data Validation

Ensuring the accuracy and reliability of collected data is essential. The following validation measures will be implemented:

- **Consistency Checks**: Survey responses will be examined for logical consistency, such as verifying that reported housing costs align with stated income levels.
- **Duplicate Detection**: Measures will be in place to identify and remove duplicate responses, using unique identifiers like address or phone number.
- **Non-Response Analysis**: Patterns of non-response will be analyzed to assess potential biases, and weighting adjustments will be applied to account for underrepresented groups.
- **Data Imputation**: For missing or incomplete responses, statistical imputation methods will be used to estimate values, maintaining the integrity of the data set.

These validation techniques are consistent with standard practices in housing surveys, ensuring the collection of high-quality data suitable for robust analysis.

B.6 Conclusion

By integrating comprehensive sampling strategies, inclusive recruitment methods, and rigorous data validation processes, this idealized survey methodology aims to provide a nuanced understanding of housing affordability in the United States. Drawing on established practices from surveys like those conducted by the CMHC, this approach addresses the limitations of existing datasets and offers valuable insights for policymakers and researchers seeking to address housing affordability challenges. While most measures of housing affordability focus on the quantitative factors, we aim to create a balance and bring in the essential qualitative factors needed to really understand how an individual or a household feels about their financial situation and housing affordability. For example, we plan to include questions about whether

respondents feel comfortable meeting their mortgage or rent payments, and if they feel secure in their ability to handle unexpected financial shocks, such as medical bills or job loss.

The survey can be found at this link: https://forms.gle/xLyZsGiEEDcwPnH66

B.7 Copy of the Survey

Housing Affordability Survey 2024

Introduction: Thank you for participating in the Housing Affordability Survey. This survey aims to understand how individuals and households perceive and experience housing affordability in the United States. Your responses will help in assessing the financial and social factors contributing to housing affordability challenges.

Participation in this survey is entirely voluntary, and you can choose to withdraw at any time without any consequences. The information collected will be used solely for research purposes. Your responses will remain confidential and will only be reported in aggregate form.

Contact Information: For questions regarding this survey or the methodology used, please reach out to:

- **Krishna Kumar**, University of Toronto. Email: krishna.kumar@mail.utoronto.ca

Section 1: Housing Situation

1. **What is your current housing situation?**

- `Own a home with a mortgage`
- `Own a home without a mortgage`
- `Rent a home/apartment`
- `Live with family/friends without paying rent`
- `Other (please specify): \_\_\_\_\_`

2. **How much of your monthly income goes toward housing costs (mortgage/rent, utilities, etc.)?**

- `Less than 20%`
- `20%-30%`

- 30%-50%

- More than 50%

3. **Do you consider your housing costs to be affordable?**

- Yes

- No

- Other: _____

4. **How would you describe the stability of your current housing situation?**

    - Very stable
    - Somewhat stable
    - Somewhat unstable
    - Very unstable

Section 2: Financial and Social Factors

5. **Have you experienced difficulty paying for housing in the past year?**

- Yes

- No

- Other: _____

6. **If you answered "Yes" to the previous question, what was the primary reason?**

- Job loss or reduced income

- Increased housing costs

- Unexpected expenses

- Other: _____

7. **Please respond according to the following statement: My current income adequately covers my housing expenses.**

- Strongly agree
- Somewhat agree
- Neutral
- Somewhat disagree
- Strongly disagree

8. **How often do you save money after covering all your monthly expenses, including housing?**

- Every month
- Most months
- Rarely
- Never

9. **If you own a home, have you refinanced or attempted to refinance your mortgage in the past year?**

- Yes
- No
- Not applicable

Section 3: Perceptions of Housing Affordability

10. **On a scale of 1 to 5, how would you rate the overall affordability of housing in your area?**

- 1 (Not affordable at all)
- 2
- 3
- 4
- 5 (Very affordable)

11. **Do you feel your local government is addressing housing affordability issues effectively?**

- Yes
- No
- Unsure

12. **Would you say that you can comfortably afford your housing?**

- Yes
- No
- Other: _____

13. **What are the biggest challenges to housing affordability in your opinion? (Select up to 3)**

- High property prices
- Rising rents
- Insufficient housing supply
- Low wages
- Lack of government support
- Other: _____

14. **How much does housing quality (size, condition, location) factor into your perception of affordability?**

    - Not at all
    - A little
    - Somewhat
    - A great deal

Section 4: Demographics

15. **What is your age range? (years)**

    - 18-24
    - 25-34
    - 35-44
    - 45-54
    - 55-64
    - 65 or older
    - Prefer not to say

16. **What gender do you identify as?**

    - Male
    - Female
    - Non-binary
    - Prefer not to say

17. **Which of the following best describes your race or ethnicity?**

- `White or Caucasian`
- `Black or African America`
- `Hispanic or Latino of any race`
- `Asian`
- `Other or multiple races`

18. **What is the highest level of education you have completed?**

    - Less than high school
    - High school diploma or GED
    - Some college, no degree
    - Associate degree
    - Bachelor's degree
    - Graduate or professional degree

- Prefer not to say

19. **What is your current employment status?**

    - Employed full-time
    - Employed part-time
    - Self-employed
    - Unemployed
    - Student
    - Retired
    - Prefer not to say

End Message: Thank you for completing this survey! Your responses are invaluable in understanding housing affordability challenges across the United States.

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.

"Average Hourly Earnings of All Employees, Total Private_2024." 2024. https://fred.stlouisfed.org/series/CES0500000003.

Boysel, Sam, and Davis Vaughan. 2021. *Fredr: An r Client for the 'FRED' API*. https://CRAN.R-project.org/package=fredr.

"Consumer Price Index for All Urban Consumers: All Items in u.s. City Average_2024." 2024. https://fred.stlouisfed.org/series/CPIAUCSL.

"Federal Funds Effective Rate_2024." 2024. https://fred.stlouisfed.org/series/FEDFUNDS.

Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40 (3): 1–25. https://www.jstatsoft.org/v40/i03/.

Iannone, Richard. 2016. *DiagrammeRsvg: Export DiagrammeR Graphviz Graphs as SVG*. https://CRAN.R-project.org/package=DiagrammeRsvg.

Iannone, Richard, and Olivier Roy. 2024. *DiagrammeR: Graph/Network Visualization*. https://CRAN.R-project.org/package=DiagrammeR.

"Median Sales Price for New Houses Sold in the United States_2024." 2024. https://fred.stlouisfed.org/series/MSPNHSUS.

"Monthly Supply of New Houses in the United States_2024." 2024. https://fred.stlouisfed.org/series/MSACSR.

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

Nguyen, Joseph. 2023. "Is There a Correlation Between Inflation and Home Prices?" *Investopedia*. https://www.investopedia.com/ask/answers/correlation-inflation-houses.asp.

Online, Et. 2024. "US Housing Crisis: Why Is America Facing Shortage of Homes and Will This Crunch Revive Inflation?" *The Economic Times*, October. https://economictimes.indiatimes.com/news/international/global-trends/us-housing-crisis-why-is-america-facing-shortage-of-homes-and-will-this-crunch-revive-inflation/articleshow/114457329.cms?from=mdr.

Ooms, Jeroen. 2024. *Rsvg: Render SVG Images into PDF, PNG, (Encapsulated) PostScript, or Bitmap Arrays*. https://CRAN.R-project.org/package=rsvg.

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

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

"Social and Affordable Housing Survey — Rental Structures Methodology_2023." 2023. https://www.cmhc-schl.gc.ca/professionals/housing-markets-data-and-research/housing-research/surveys/methods/methodology-social-affordable-housing-survey-rental-structures.

Wickham, Hadley. 2011. "Testthat: Get Started with Testing." *The R Journal* 3: 5–10. https://journal.r-project.org/archive/2011-1/RJournal_2011-1_Wickham.pdf.

———. 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, Jim Hester, and Jennifer Bryan. 2024. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.