My title*

My subtitle if needed

Krishna Kumar

November 23, 2024

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

1 Introduction

Overview paragraph

Estimand paragraph

Results paragraph

Why it matters paragraph

Telegraphing paragraph: The remainder of this paper is structured as follows. Section 2….

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` **(CITE)** package and loaded using the same library. The data were cleaned and analysed using the statistical programming language `R`. Simulating, cleaning, testing, and graphing were done with the help of the following packages: `Tidyverse`, `dplyr`, `lubridate`, `testthat`, `readr`, `arrow`, `ggplot2`, `here`, **(CITE + ADD ANY MORE YOU USED HERE) this**

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

The Federal Research Economic Data provides information 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. **(CITE 1) FROM REFERENCES)** Studying this data is essential in assessing the economic factors affecting housing affordability as well as the financial wellbeing of individuals. The data sets were all downloaded separately using the FRED API and then combined during the data cleaning.

2.2 Measurement (LEFT TO DO)

For each variable:

- `CPI`:
- `Interest Rate`:
- `Average Wage`:
- `House Price`:
- `House Supply`:

2.3 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 _____ and to help with year over year change analysis of the house prices and wage rates.

2.4 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.

2.5 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. **(CITE 2) HERE FROM REFERENCES)**
- **`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. **(CITE 3) HERE FROM REFERENCES)**
- **`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. **(CITE 4) FROM REFERENCES)**
- **`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. **(CITE 5) FROM REFERENCES)**

- **`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. **(CITE 6) FROM REFERENCES)**
- **`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.

2.5.1 Consumer Price Index Trends Over Time

Figure **(CROSS REFERENCE HERE)** 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 homeownership, 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. **(CITE 7) HERE FROM REFERENCES)**

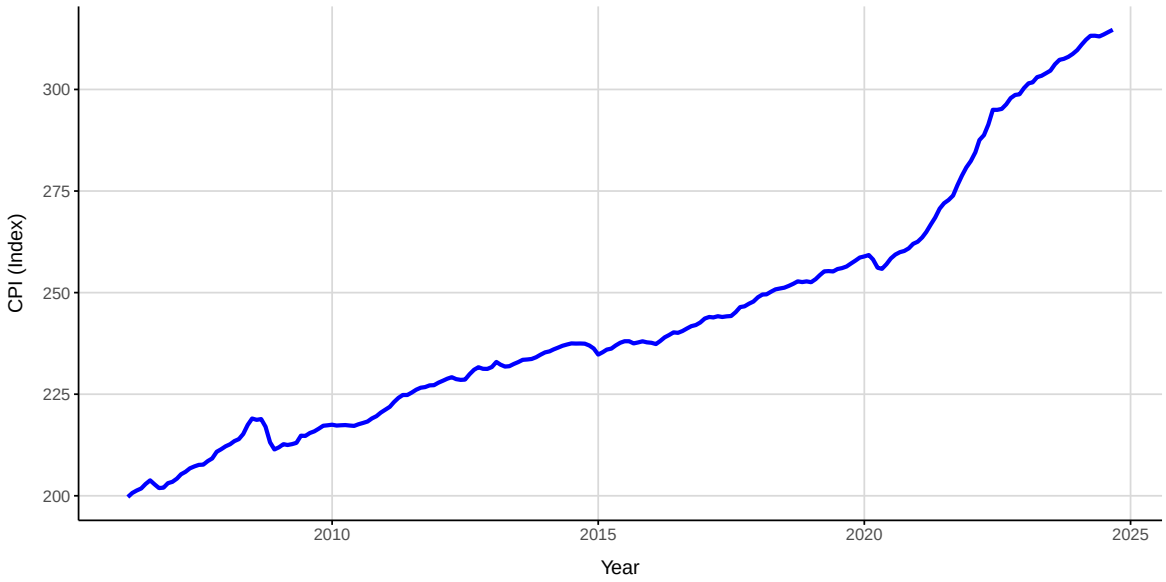

Figure 1: Consumer Price Index (CPI) Over Time

2.5.2 Interest Rates Trends Over Time

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

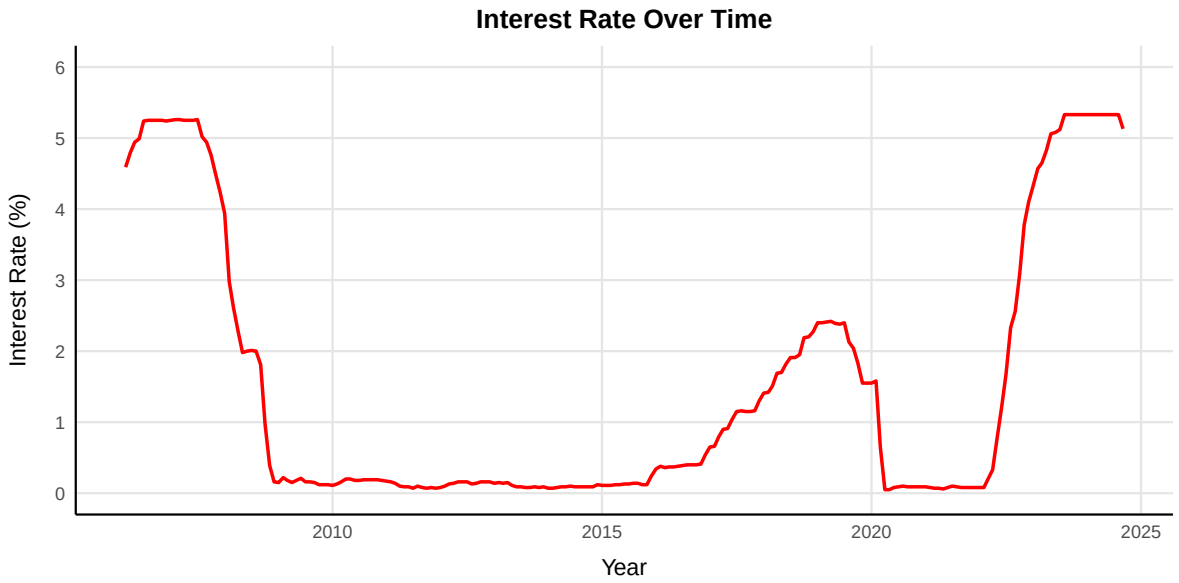

Figure 2: Federal Funds Effective Interest Rate Over Time

2.5.3 Housing Supply Over Time

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.

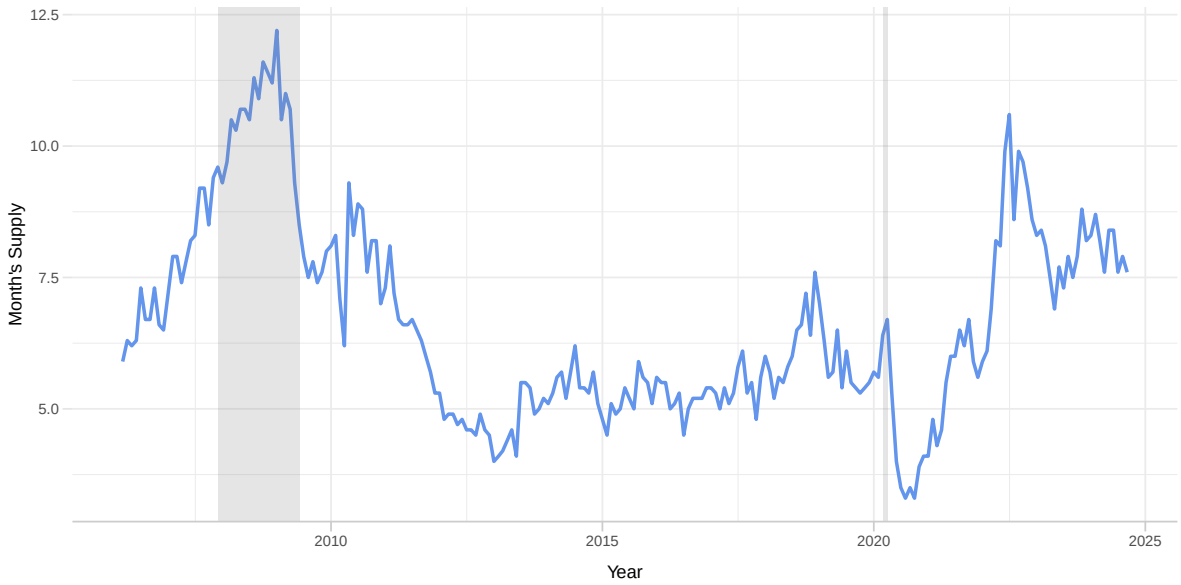

Figure 3: Housing Supply Over Time

2.5.4 Wage Growth vs House Prices

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.

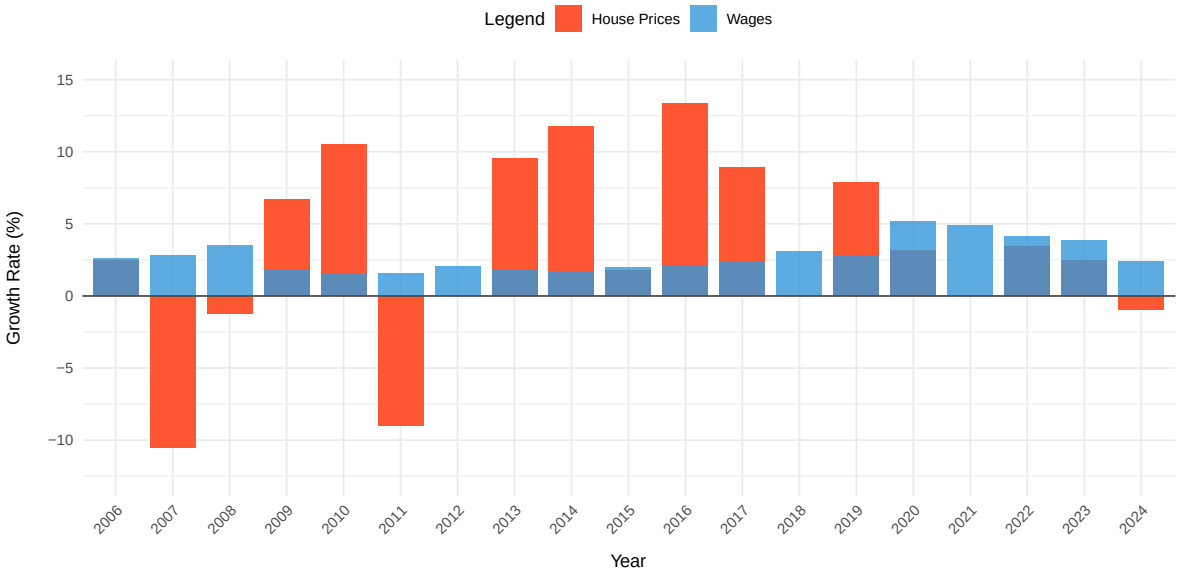

Figure 4: Growth Rates of House Prices and Wages

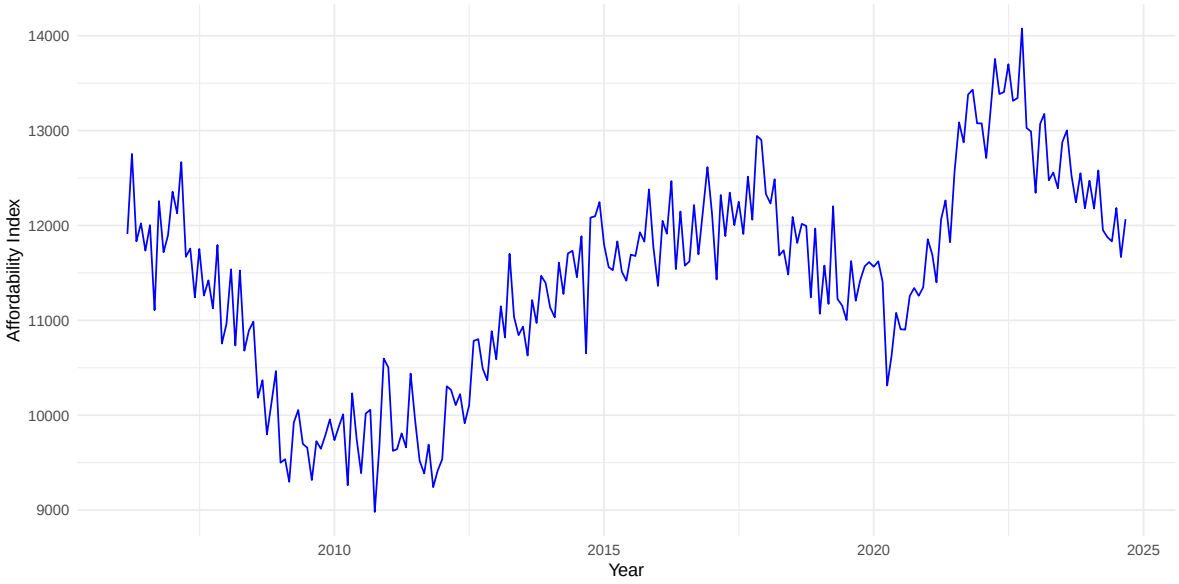

Figure 5: Growth Rates of House Prices and Wages

2.5.5 Affordability Index (CLEAN THIS UP)

3 Model

```r
# Load necessary library
library(tidyverse)

# Fit a simple multiple linear regression model
model <- lm(
  formula = `Affordability Index` ~ `Interest Rate (%)` + `CPI (Index)`,
  data = analysis_data
)

# Summarize the model
summary(model)
```

```
Call:
lm(formula = `Affordability Index` ~ `Interest Rate (%)` + `CPI (Index)`,
    data = analysis_data)

Residuals:
     Min       1Q   Median       3Q      Max
-1761.29  -648.94   -39.14   649.02  1704.67

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)        6464.677    452.480  14.287  < 2e-16 ***
`Interest Rate (%)`  144.409     29.367   4.917 1.72e-06 ***
`CPI (Index)`         19.398      1.878  10.331  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 816 on 220 degrees of freedom
Multiple R-squared:  0.4325,    Adjusted R-squared:  0.4273
F-statistic: 83.82 on 2 and 220 DF,  p-value: < 2.2e-16
```

ADD WORDS HERE.

The goal of our modelling strategy is twofold. Firstly,...

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

3.1 Model set-up

We aim to model the **Affordability Index**, which represents the number of hours of work required to afford a house, 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

ADD WORDS HERE.

We expect a positive relationship between the size of the wings and time spent aloft. In particular…

We can use maths by including latex between dollar signs, for instance θ.

ALSO ADD A DAG HERE.

Table 1: Explanatory models of flight time based on wing width and wing length

	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.

TALK ABOUT RESULTS AND THEIR MEANING/IMPLICATIONS HERE.

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

Please don't use these as sub-heading labels - change them to be what your point actually is.

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

In **?@fig-ppcheckandposteriorvsprior-1** we implement a posterior predictive check. This shows…

In **?@fig-ppcheckandposteriorvsprior-2** we compare the posterior with the prior. This shows…

Examining how the model fits, and is affected
by, the data

B.2 Diagnostics

?@fig-stanareyouokay-1 is a trace plot. It shows… This suggests…

?@fig-stanareyouokay-2 is a Rhat plot. It shows… This suggests…

Checking the convergence of the MCMC algo-
rithm

C References