**Cornell University**

**Covid-19 Forecasting**

**Spring 2020**

**Jacob Shusko, John Miller, and Catherine Appleby**

**May 12th, 2020**

# Abstract

As Covid-19 has spread across the United States, there have been a variety of non-pharmaceutical interventions (NPIs) put into effect by counties to reduce case load and prevent the spread of infection. As states and counties begin to re-open businesses and public spaces, it is important to know which counties will be at risk for large outbreaks and to know an effective timeline for instituting effective NPIs to curb these outbreaks. Based on the type of NPI in use, the time between [first infection? Covid-19 pandemic status?] and its start date, and demographics, we forecast the risk level of a Covid-19 outbreak for counties across the United States. Using Covid-19 testing and death data by county and date joined with NPI and demographic data to form a time series dataset, we fit an ARIMA model to predict risk levels. [Issues found, disovery/quick conclusion]

**Contents**

[Abstract 2](#_Toc39651389)

[Introduction 4](#_Toc39651390)

[Preprocessing Data 5](#_Toc39651391)

[Methods 6](#_Toc39651392)

[Results 7](#_Toc39651393)

[Conclusion 8](#_Toc39651394)

# Introduction

We have solved the coronavirus by being smort and interrigent. ARIMA is life.

Catherine will copy and paste the introduction here after she’s had some coffee. <- this is the new text btw…

# Preprocessing Data

To create our master data set, we merged together three data sets on the county level: a Census data set containing demographic information for each county, a NPI data set provided by Keystone Strategy, and a Covid-19 deaths data set provided by John Hopkins.

# Methods

# Results

# Conclusion

# References