
Ali Ahmed
Ph.D. Candidate Management Information Systems
Department of Operations & Information Systems
Manning School of Business
University of Massachusetts Lowell
72 University Ave, Lowell, MA 01854 USA
Cell: 832-893-0837
Email: ali_ahmed@student.uml.edu

Abstract

The importance of testing and surveillance of an infectious disease cannot be underestimated. The testing is the first step to detect an infectious disease, and mass testing can slow or mitigate the spread of an infectious disease. Despite overwhelming evidence and the importance of testing discussed in the literature, there have been claims that “more COVID-19 testing creates more cases”. Therefore, there is a need to study whether massive testing is the reason for detecting more positive COVID-19 cases. In this research, we used a dataset from the U.S. Department of Health & Human Services and empirically showed that by increasing the COVID-19 testing in the U.S., the spread of the COVID-19 decreased significantly. Our results indicate a negative relationship between the number of positive cases and the number of tests performed in the past months. The large-scale testing may have helped identify positive and asymptomatic cases early in the course of illness, which enabled individuals to isolate themselves, thus reducing the chances of spreading the diseases and slowing the spread of the pandemic.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Introduction

Detection and surveillance of an infectious disease is the first step in controlling a pandemic\(^1\). However, not much attention has been given to understand the role of testing in mitigating an infectious disease’s spread\(^2\). In this research, we present empirical evidence of how COVID-19 testing effects the spread of the disease and the number of positive cases in the future. We propose increasing the COVID-19 testing capacity in a given population may help identify asymptomatic and early cases, resulting in deterrence and decreased positive cases in the future. Specifically, we have analyzed a relationship between the number of positive cases reported in a month and the number of tests performed in the past months.

Methods

We used a publicly available dataset released by the U.S. Department of Health & Human Services\(^3\). The dataset includes the COVID-19 laboratory test [Polymerase chain reaction (PCR)] results from over 1,000 U.S. laboratories and testing locations, including commercial and reference laboratories, public health laboratories, hospital laboratories, and other testing locations. The dataset contains the daily number of COVID-19 positive, negative, and inconclusive test results reported by each U.S. state’s health department. Using this dataset, we created a state-month panel data that includes the number of cumulative positive, negative, and inconclusive test results in each U.S. state for each month, ranging from March to November 2020.

To test the effect of COVID-19 testing on the number of positive cases in the coming months, we estimated two fixed-effects multiple linear regression models. Equation 1 analyzes the relationship between the number of tests conducted in the last month and the number of positive cases in the focal month. The main independent variable is the \( \text{LogTesting}_{t-1} \). The Equation 2
analyzes the relationship between the cumulative number of tests performed
\((CumulativeTest_{i,k})\) in all previous months and the number of positive cases reported in the
focal month. The dependent variable for both equations is the log-transformed number of
positive cases reported in the focal month \(t\) by the U.S. State \(i\).

\[
LogPositiveCases_{i,t} = \alpha_0 + \alpha_1 LogPositiveCases_{i,t-1} + \alpha_2 LogTesting_{i,t} + \alpha_3 LogTesting_{i,t-1} + State_i + Month_t + \epsilon_{i,t} \tag{1}
\]

\[
LogPositiveCases_{i,t} = \beta_0 + \beta_1 LogPositiveCases_{i,t-1} + \beta_2 LogTesting_{i,t} + \beta_3 CumulativeTest_{i,k} + State_i + Month_t + \epsilon_{i,t} \tag{2}
\]

where \(CumulativeTest_{i,k} = \text{Log}(\Sigma_{k=1}^{t-1} Testing_{i,k})\) \(\tag{3}\)

In Equation 1 – 3, \(i\) represents the 50 U.S. States, \(t\) indicates the month dummies on which the
test results were conducted, ranging from April – November. The subscript \(k\) represents the
lagged months, ranging from 1 – 10. The variable \(State_i\) is the fixed-effects controlling for the
unobserved time-invariant characteristics of a U.S. state \(i\). The variable \(Month_t\) is the time-
dummy variable controlling for variations in disease spread from month to month. Finally, \(\epsilon_{i,t-k}\)
is the idiosyncratic error of the models.

**Results**

Our results suggest a negative relationship between the number of positive cases and the number
of tests performed in the past months. In Table 1, Column 1 and 2 shows the estimated models
for Equation 1 and 2, respectively. Column 1 shows the negative effect of increasing the number
of tests performed in the last month \((LogTesting_{i,t-1})\) on the number of cases on the focal month
within a State. Column 2 shows a similar pattern of decrease in positive cases by an increase in
the number of tests performed in all the previous months (i.e. \( \text{Cumulative}_{t,t-1} \)). The overall relationship suggests that having more tests in previous months reduces the number of positive cases in the coming months. Figure 1 shows the predicted relationship for both the models.

Table 1: Effect of Testing on Positive Cases in Future Months

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \text{LogPositive}_{i,t} )</td>
<td>( \text{LogPositive}_{i,t} )</td>
</tr>
<tr>
<td>( \text{LogPositive}_{i,t-1} )</td>
<td>0.6337*** (0.04875)</td>
<td>0.5962*** (0.04642)</td>
</tr>
<tr>
<td>( \text{LogTesting}_{i,t} )</td>
<td>1.1727*** (0.1146)</td>
<td>1.0546*** (0.1052)</td>
</tr>
<tr>
<td>( \text{LogTesting}_{i,t-1} )</td>
<td>-0.6064*** (0.1096)</td>
<td>-</td>
</tr>
<tr>
<td>( \text{CumulativeTest}_{i,t-1} )</td>
<td>-</td>
<td>-0.5595*** (0.09651)</td>
</tr>
<tr>
<td>( \text{Month}_{t} = \text{April} )</td>
<td>0.4723** (0.1897)</td>
<td>-0.8573** (0.3506)</td>
</tr>
<tr>
<td>( \text{Month}_{t} = \text{May} )</td>
<td>-0.4543*** (0.1353)</td>
<td>-1.5794*** (0.2678)</td>
</tr>
<tr>
<td>( \text{Month}_{t} = \text{June} )</td>
<td>-0.2888*** (0.1059)</td>
<td>-1.2548*** (0.1901)</td>
</tr>
<tr>
<td>( \text{Month}_{t} = \text{July} )</td>
<td>0.03250 (0.07688)</td>
<td>-0.7411*** (0.1551)</td>
</tr>
<tr>
<td>( \text{Month}_{t} = \text{August} )</td>
<td>-0.2606*** (0.07552)</td>
<td>-0.9515*** (0.1123)</td>
</tr>
<tr>
<td>( \text{Month}_{t} = \text{September} )</td>
<td>-0.3222*** (0.06791)</td>
<td>-0.8028*** (0.09657)</td>
</tr>
<tr>
<td>( \text{Month}_{t} = \text{October} )</td>
<td>-</td>
<td>-0.3574*** (0.1012)</td>
</tr>
<tr>
<td>( \text{Month}_{t} = \text{November} )</td>
<td>0.4031*** (0.1026)</td>
<td>-</td>
</tr>
<tr>
<td>( \text{Constant} )</td>
<td>-2.9847*** (0.7316)</td>
<td>-0.9512 (0.7717)</td>
</tr>
<tr>
<td>State Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>430</td>
<td>430</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.939</td>
<td>0.937</td>
</tr>
</tbody>
</table>

** p<0.05, *** p<0.01

Discussion

Our research explains the relationship between COVID-19 testing and positive cases. We provide empirical evidence on the importance of mass testing and its effects on controlling the
pandemic. These analyses also refute the misperception that more tests yield more positive cases. We have shown that increasing the number of the tests yield fewer cases in the future—testing and surveillance help identify asymptomatic cases, which create deterrence and decrease the disease’s spread\(^2\). In short, mass testing mitigates the spread and yields less positive cases in the coming months.

Figure 1: Relationship between the Number of Positive Cases and Tests in Previous Months
References

