An integration of SIRD-model and crude oil pricing under COVID-19

Yuhan Hu and Hoi Ying Wong

July 15, 2021

Abstract

Under COVID-19 pandemic, social distancing is strictly controlled to reduce the spread of the infectious disease. The demand of crude oil drops a lot with the sharp decline of mobility at the beginning of the pandemic. Sudden decrease of demand and limited storage space cause negative oil prices. The commonly used oil pricing models behave unsatisfying under the infectious disease. This paper investigates how the pandemic affects the crude oil prices. We first formulate the relationship between infection rate and social distancing, then we integrate COVID-19 data into oil pricing. By applying this framework to US data, we use the model to value future contracts and contrast its behavior with other classic models. The results show that the COVID-19 statistics can significantly enhance the behavior of valuation for crude oil futures and the pandemic has long-lasting effects for oil market.

Keywords: Crude Oil: Futures, COVID-19, Stochastic SIRD Model, Model Implementation

1 Introduction

The coronavirus disease 2019 (COVID-19) is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and first identified in Dec 2019, Wuhan, China. It globally broke out from Feb 2020 and was announced as a pandemic by WHO on 11 March 2020. The worldwide infection has caused millions of deaths, lock-down between countries, stock market crash, and huge volatility of oil prices. Policymakers kept a strict social distancing to avoid further spread due to the short of medical facilities and lack of full vaccination. At that time, citizens were enforced to stay at home, plenty of flights suspended and even some cities were locked down. As a consequence of the extremely crash in demand, a historic decrease of WTI crude oil price occurred on April 20, which dropped by almost 300% and touched -37 dollar per barrel [10].

Bachelier is the first to model stochastic change in financial instrument prices using Brownian Motion [7]. However, Bachelier model is barely used in commodity valuation. Black [9] introduces Black model in 1976, which is widely used to price commodity option among practitioners. Gibson and Schwartz developed Gibson and Schwartz two-factor model in 1990. This model [12] first introduces instantaneous convenience yield into oil pricing, and assume it follows a Ornsteion-Uhlenbeck process. Schwartz (1997) then extended the two-factor model to include stochastic interest rates, and assume the interest rates follow a mean-reverting process [23]. However, the commonly used classic models fail to price oil derivatives with negative spot price. The Chicago Mercantile Exchange and Intercontinental Exchange announced changing Black model to Bachelier model to price oil futures options [4, 5].

Some papers evaluate the relationship between the coronavirus disease and oil prices. [21] investigates how the COVID-19 pandemic and oil price news affect oil prices. [6] shows that the coronavirus new infections have a marginal negative influence on the oil price in the long term. Using some fractional integration techniques, Gil-Alana and Manuel [13] show that oil price series is mean reverting, which indicates COVID-19 will be a transitory shock with long-lasting effects. And some papers investigate the relationship between crude oil and stock markets under COVID-19 [19, 22]. To the best of our knowledge, no research has developed an exact model to describe how the COVID-19 pandemic affects oil prices, which motivates us to work on that.

Our paper uses SIRD model to show the current epidemiological situation. Then we propose a new framework combining mobility and COVID-19 statistics into crude oil pricing. To see how the public mobility affects the oil prices, we raise two models. One includes mobility, while the other doesn't. Finally, we focus on the prediction of future prices. The empirical results of this paper show that the performance of the model with mobility and COVID-19 statistics is superior to other classic pricing models like Bachelier model and mean reversion model. Although the behavior of other models are worse in pricing longer term future contracts, our model is still satisfying and robust, which supports the long-lasting influence of the pandemic in crude oil market [13]. The outstanding performance of the model integrated with mobility in pricing with different term also supports that the significance of social distancing even with the relative high vaccine inoculation rate [16].

The remainder of the article is organized as follows. In section 2, we incorporate social distancing into stochastic epidemic model, and then implement this framework to US COVID-19 data. In section 3, we derive two models to price crude oil, one includes COVID-19 statistics and mobility, the other excludes mobility. In section 4, we estimate the parameters of the joint stochastic pricing process introduced in section 3 over the coronavirus infection period. Then we apply the models to value WTI oil futures and contrast their performances with Bachelier model and mean reversion model. In section 5, we conclude the paper.

2 Stochastic Modeling of COVID-19

Kermack and Mckendrick [18] created SIR model in early 20th century, which is the deterministic mathematical framework for epidemic diseases. Many derivatives of the basic model contain more complicated compartments, such as those include births, deaths or immunity [8]. Varieties of stochastic models are developed to forecast and control COVID-19. [11] uses worldwide infection and death data to estimate a standard epidemiological model of COVID-19. [17] models the spread by considering additional individual action, and [15, 20] investigate how the government's control policies affect the transmission. [14] builds a SIDARTHE model and suggests that social-distancing measures and population-wide testing would be very necessary and effective ways to stop further infections at the early stage. We model the COVID-19 process with the SIRD model instead of other more complicated models after considering the availability of data and simplicity of framework to price crude oil in section 3.

2.1 SIRD Model and Mobility

In the SIRD model, the whole population is assigned to compartments with labels, S, I, R, D, which represents the fractions of susceptible, infectious, recovered and deceased individuals due to the epidemic, respectively. Therefore, we only require the data of total cases, active cases and deaths of the coronavirus disease downloaded from Worldometer.info; see [3] .

The SIRD model uses the following system of differential equations:

$$\begin{cases} dI_t = (\beta S_t - \delta - \gamma) I_t dt \\ dD_t = \delta I_t dt \\ dR_t = \gamma I_t dt \\ dS_t = \beta S_t I_t dt \end{cases} \tag{1}$$

where β, γ, δ are the rates of infection, recovery and mortality, respectively. And S_t, I_t, R_t, and D_t are the susceptible, infectious, recovered and deceased rates, calculated by number of labeled individuals divided by total population, with summation equal to 1:

$$I_t + R_t + D_t + S_t = 1 \tag{2}$$

When the outbreak of COVID-19 occurred in the initial phase, many countries adopted strict quarantine policies, some even locked down the entire city to keep the spread of the infectious disease [24]. The vaccination for COVID-19 was available for the public from December 2020, and 22.6% people in the world has received at least one dose of the COVID-19 vaccine [1]. [16] shows that only the single vaccination process cannot prevent COVID-19 resurgence but the integration of vaccination and social distancing can achieve that. Thus, it's reasonable to postulate the infection rate of the susceptible population is dependent of mobility.

We assume that recovered people will get antibody against second infection, and the vaccination effect to the pandemic situation has already been incorporated into the infection and death data. The infection rate of the susceptible population, βS_t, is defined as B_t in this paper. We put forward two ways to predict COVID-19 statistics. One is that B_t is linearly associated with the moving average of mobility $\bar{\alpha}_t$, as the social distancing control can reduce transmission to susceptible population:

$$B_t = a_0 + a_1 \bar{\alpha}_t + \epsilon_t, \qquad \epsilon_t \sim \mathbf{N}\big(0, \sigma_\epsilon^2\big) \tag{3}$$

where α_t represents the mobility, ϵ_t represents the residual errors in (3). From SIRD framework (1), I_t can be represented by the function of B_t:

$$I_t = I_{t_0} e^{\int_{t_0}^{t} B_s - \delta - \gamma ds} \tag{4}$$

where t_0 is the initial time. Thereafter, the COVID-19 statistics can all be expressed by functions of B_t:

$$\begin{cases} I_t = I_{t_0} e^{\int_{t_0}^{t} B_s - \delta - \gamma ds} \\ D_t = D_{t_0} + \delta I_{t_0} \int_{t_0}^{t} e^{\int_0^s B_\tau - \delta - \gamma d\tau} ds \\ R_t = R_{t_0} + \gamma I_{t_0} \int_{t_0}^{t} e^{\int_{t_0}^s B_\tau - \delta - \gamma d\tau} ds \end{cases} \tag{5}$$

The other way to estimate COVID-19 statistics is to assume β as a constant, and I_t, D_t, R_t can be estimated by $I_{t-\Delta t}$, $R_{t-\Delta t}$, $D_{t-\Delta t}$:

$$\begin{cases} S_t = S_{t-\Delta t} - \beta I_t S_t \Delta t \\ D_t = D_{t-\Delta t} + \delta I_t \Delta t \\ R_t = R_{t-\Delta t} + \gamma I_t \Delta t \\ I_t = 1 - S_t - R_t - D_t \end{cases} \tag{6}$$

where t_0 is the initial time and $t - \Delta t \geq t_0$.

2.2 US COVID-19 Data

Data of total cases, active cases and deaths of the coronavirus disease are downloaded from Worldometer.info; see [3]. B_t is calculated by $\log(\frac{I_{t+\Delta t}}{I_t}) + \delta + r$ with collected data in the period from 2020-Feb-20 to 2021-Mar-22. We first regress B_t on the 5-day moving average mobility $\bar{\alpha}_t$ and conduct the Shapiro-Wilk normality tests with residuals ϵ_t. The linear relationship between the infection rate to susceptible population and moving average mobility is only 0.05, and the residuals are not normally distributed either. We also collect the mobility indices from Google Mobility [2] across six different categories of places: retail and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential. And then regress B_t on the 5-day moving average mobility indices, which is written as $\bar{\alpha}_t^M = \begin{bmatrix} \bar{\alpha}_t^{RR} & \bar{\alpha}_t^{GP} & \bar{\alpha}_t^{PA} & \bar{\alpha}_t^{TS} & \bar{\alpha}_t^{WP} & \bar{\alpha}_t^{RE} \end{bmatrix}^T$. As shown in table 2, the coefficients of determination is 0.57 with multiple mobility moving averages, which supports the linear association between B_t and public mobility, while the residuals of the regression are not obey Gaussian distribution.

Table 1: Regression table of fitting B_t with single mobility moving average

a_0	a_1	σ_ϵ	R^2	p-value
0.1345	0.2648	0.1222	0.0566	<2.2e-16

OLS model: $\boldsymbol{B_t = a_0 + a_1\bar{\alpha}_t + \epsilon_t}$

The last column reports the p-value of the Shaprio-Wilk normality test of the residuals. σ_ϵ is the residual's standard error.

Table 2: Regression table of fitting B_t with multiple mobility moving averages

a'_0	a'_1	a'_2	a'_3	a'_4	a'_5	a'_6	σ'_ϵ	R^2	p-value
0.3470	-0.5778	0.6556	-0.1425	1.7456	0.2487	3.1825	0.0829	0.5738	<2.2e-16

OLS model: $B_t = a'_0 + a'_1\bar{\alpha}_t^{RR} + a'_2\bar{\alpha}_t^{GP} + a'_3\bar{\alpha}_t^{PA} + a'_4\bar{\alpha}_t^{TS} + a'_5\bar{\alpha}_t^{WP} + a'_6\bar{\alpha}_t^{RE} + \epsilon'_t$

The last column reports the p-value of the Shaprio-Wilk normality test of the residuals. σ'_ϵ is the residual's standard error.

3 Price Crude Oil under Coronavirus Disease

We propose two oil pricing frameworks under coronavirus disease. One suggests crude oil spot price is linearly associated with all the lagged COVID-19 statistics $I_{t-\Delta t}$, $R_{t-\Delta t}$, $D_{t-\Delta t}$, mobility $\alpha_{t-\Delta t}$ and the lagged oil price $P_{t-\Delta t}$, named SIRDMP model. The other excludes mobility, named SIRDP model. We use mobility α_t instead of its moving average as the market has immediate reaction to public mobility change.

We postulate the regression models:

$$P_t = K_0 + K_P P_{t-\Delta t} + K_\alpha \alpha_{t-\Delta t} + K_I I_{t-\Delta t} + K_R R_{t-\Delta t} + K_D D_{t-\Delta t} + e_t \tag{7}$$

$$P_t = k_0 + k_P P_{t-\Delta t} + k_I I_{t-\Delta t} + k_R R_{t-\Delta t} + k_D D_{t-\Delta t} + e'_t \tag{8}$$

The regression model can be written as follow if we use multiple mobility indices:

$$P_t = \hat{K}_0 + \hat{K}_P P_{t-\Delta t} + \hat{K}_\alpha^T \alpha_{t-\Delta t}^M + \hat{K}_I I_{t-\Delta t} + \hat{K}_R R_{t-\Delta t} + \hat{K}_D D_{t-\Delta t} + \hat{e}_t \tag{9}$$

for $t = t_0,\ t_1,\ t_2, ...,\ T$. K_0, K_P, K_α, K_ξ, K_I, K_R, K_D are regression coefficients and e_t is the error term for equation (7). k_0, k_P, k_ξ, k_I, k_R, k_D are regression coefficients and e'_t is the error term for equation (8). $\hat{K}_0$, $\hat{K}_P$, $\hat{K}_\alpha^T \in \mathbb{R}^6$, $\hat{K}_\xi$, $\hat{K}_I$, $\hat{K}_R$, $\hat{K}_D$ are regression coefficients and $\hat{e}_t$ is the error term for equation (9). We exhibit the predicted oil prices from (7) and (8) by daily data to show the ability to derive negative values of oil price of our framework. CME and ICE announced changing Black-schole model to Bachelier model to price oil futures options [4, 5], as Black-schole model cannot generate negative values with initial positive prices. Our framework conquers this shortcoming and add more information about the carnivorous pandemic, which raises the accuracy in pricing futures as shown in section 4.

Figure 1: Predicted values of oil price by equation (7) and equation (8) using weekly data and daily data, respectively.

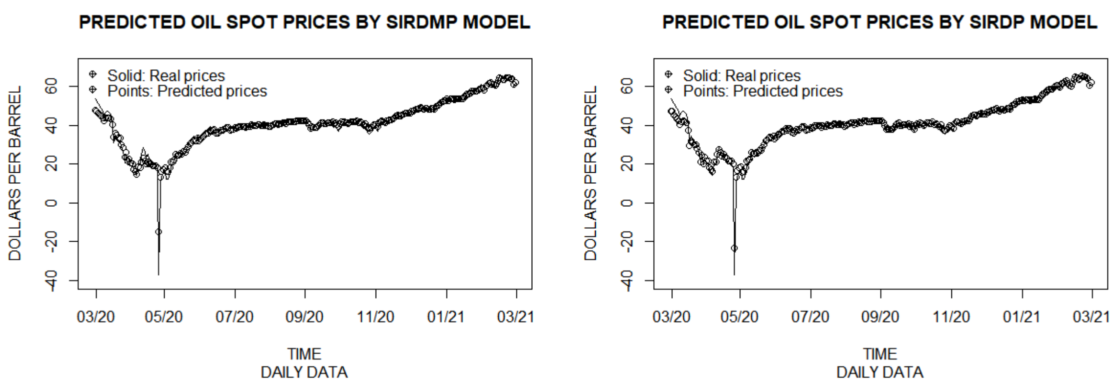

Table 3, 4 and 5 summarize the regression results. The coefficients of determination of oil spot prices on all the above factors are above 0.88 without containing mobility, which shows that the COVID-19 statistics, and spot prices have high explanatory and forecast power for oil pricing. R^2 increase a little after including the effect of mobility. As we can see in table 3 and 4, relative weight of lagged spot prices is largest, and COVID-19 statistics contributes almost evenly in the prediction of oil prices. Table 5 shows that the summation of all relative weights of the factors in the mobility indices is near to the relative weight of single mobility. Multicollinearity exists between the pricing factors, which requires more cautions in explaining a single regression coefficient. Nevertheless, multicollinearity won't affect the crude oil price prediction as such relationship has already been adjusted by the coefficients.

Table 3: Regression table of fitting oil spot with lagged spot prices, mobility and COVID-19 statistics.

	(Intercept)	P_{t-1}	α_{t-1}	I_{t-1}	R_{t-1}	D_{t-1}
($R^2 = 0.8937$, p-value$<$ 2.2e-16)						
coefficients	16.3389	0.5882	19.8127	216.0134	-261.8855	6624.8130
relative weight	-	30.6303	14.6652	18.1498	17.6107	18.9441
multicollinearity	-	True	False	True	True	True

P-value reported in this table is from the Shaprio-Wilk normality test of the residuals.

Table 4: Regression table of fitting oil spot with lagged spot prices and COVID-19 statistics.

	(Intercept)	P_{t-1}	I_{t-1}	R_{t-1}	D_{t-1}
($R^2 = 0.8821$, p-value$<$ 2.2e-16)					
coefficients	4.9897	0.7898	-23.5064	-29.6306	7506.3508
relative weight	-	45.0203	17.7387	17.8431	19.3978
multicollinearity	-	False	True	True	True

P-value reported in this table is from the Shaprio-Wilk normality test of the residuals.

Table 5: Regression table of fitting oil spot with lagged spot prices, mobility indices and COVID-19 statistics.

	(Intercept)	P_{t-1}	I_{t-1}	R_{t-1}	D_{t-1}	RR_{t-1}
($R^2 = 0.8978$, p-value$<$ 2.2e-16)						
coefficients	17.2810	0.5463	310.1577	-278.1259	5608.5977	-9.2171
relative weight	-	27.3997	17.8788	17.5478	18.5798	4.0981
multicollinearity	-	True	True	True	True	True
		GP_{t-1}	PA_{t-1}	TS_{t-1}	WP_{t-1}	RE_{t-1}
coefficients	-	-7.4684	6.0790	26.2483	3.1478	1.5175
relative weight	-	1.7810	2.5034	3.2326	3.0367	3.9419
multicollinearity	-	True	True	True	True	True

P-value reported in this table is from the Shaprio-Wilk normality test of the residuals.

3.1 SIRDMP Pricing Model with single mobility

To see the contribution of mobility to the prediction of COVID-19 statistics and crude oil prices, we first build an oil pricing framework with public mobility α_t, which is postulated to follow a mean-reverting process:

$$d\alpha_t = b(\theta_\alpha - \alpha_t)dt + \sigma_\alpha dW_\alpha \tag{10}$$

For the SIRDMP model, we subtract $P_{t-\Delta t}$ from equation (7) in both sides, then we can derive a PDE for oil price P_t:

$$dP_t = (\theta_P^1(t) - \lambda^1 P_t)dt + \sigma_P^1 dW_P \tag{11}$$

where

$$\theta_P^1(t) = K_0 + K_\alpha\alpha_t + K_I I_t + K_R R_t + K_D D_t \tag{12}$$

and

$$dW_P dW_\alpha = \rho_1 dt. \tag{13}$$

As shown in section 2, all the COVID-19 statistics can be written as a function of B_t, and B_t is predicted by moving average of mobility $\bar{\alpha}_t$. Thus, θ_P can be written as:

$$\begin{aligned}\theta_P^1(t) = K_0 + K_\alpha\alpha_t + K_I I_{t_0} e^{\int_{t_0}^{t} B_s - \delta - \gamma ds} + K_R(R_{t_0} + \gamma I_{t_0}\int_{t_0}^{t} e^{\int_{t_0}^{s} B_\tau - \delta - \gamma d\tau} ds) \\ + K_D(D_{t_0} + \delta I_{t_0}\int_{t_0}^{t} e^{\int_0^s B_\tau - \delta - \gamma d\tau} ds)\end{aligned} \tag{14}$$

where

$$B_t = a_0 + \hat{a}_1 \int_{t-h}^{t} \alpha_s ds + \epsilon_t \text{ with } \hat{a}_1 = \frac{a_1}{h} \tag{15}$$

and

$$\alpha_t = \alpha_{t_0} e^{-b(t-t_0)} + \theta_\alpha(1 - e^{-b(t-t_0)}) + \sigma_\alpha \int_{t_0}^{t} e^{-b(t-s)} dW_\alpha. \tag{16}$$

The spot prices of crude oil P_T has such closed-form solution:

$$P_T^1 = P_t e^{-\lambda^1(T-t)} + e^{-\lambda^1 T}\int_t^T e^{\lambda^1 s}\theta_P^1(s)ds + \sigma_P^1 \int_t^T e^{-\lambda^1(t-s)} dW_P. \tag{17}$$

3.2 SIRDMP Pricing Model with mobility indices

If we use mobility indices $\alpha_t^M = \begin{bmatrix}\alpha_t^{RR} & \alpha_t^{GP} & \alpha_t^{PA} & \alpha_t^{TS} & \alpha_t^{WP} & \alpha_t^{RE}\end{bmatrix}^T$, the joint stochastic process of P_t and α_t^M will be:

$$d\alpha_t^M = \mu_\alpha^M dt + \sigma_\alpha^M dW_\alpha^M \tag{18}$$

$$dP_t = (\theta_P^2(t) - \lambda^2 P_t)dt + \sigma_P^2 dW_P \tag{19}$$

$$dW_\alpha^M dW_P = \rho_1^M dt. \tag{20}$$

$$
\begin{aligned}
\theta_P^2(t) &= \hat{K}_0 + \hat{K}_\alpha^T \alpha_t^M + \hat{K}_I I_t + \hat{K}_R R_t + \hat{K}_D D_t \\
&= \hat{K}_0 + \hat{K}_\alpha^T \alpha_t^M + \hat{K}_I I_{t_0} e^{\int_{t_0}^t \hat{B}_s - \delta - \gamma ds} + \hat{K}_R (R_{t_0} + \gamma I_{t_0} \int_{t_0}^t e^{\int_{t_0}^s \hat{B}_\tau - \delta - \gamma d\tau} ds) \\
&+ \hat{K}_D (D_{t_0} + \delta I_{t_0} \int_{t_0}^t e^{\int_0^s \hat{B}_\tau - \delta - \gamma d\tau} ds)
\end{aligned} \tag{21}
$$

where

$$
\hat{B}_t = a_0' + \hat{a}_1^T \int_{t-h}^t \alpha_s^M ds + \epsilon_t' \quad \text{with} \quad \hat{a}_1^T = \frac{a_1^T}{h} \tag{22}
$$

where $\hat{a}_1 = \begin{bmatrix} a_1' & ... & a_6' \end{bmatrix}^T$, $\mu_\alpha^M = \begin{bmatrix} b_1(\theta_{RR} - \alpha_t^{RR}) & ... & b_6(\theta_{RE} - \alpha_t^{RE}) \end{bmatrix}^T$, $\sigma_\alpha^M = \begin{bmatrix} \sigma_{RR} & ... & \sigma_{RE} \end{bmatrix}^T$, $\rho_1^M = \begin{bmatrix} \rho_1^{RR} & ... & \rho_1^{RE} \end{bmatrix}^T$, $dW_\alpha^M = \begin{bmatrix} dW_{RR} & ... & dW_{RE} \end{bmatrix}^T$ are all column vectors with size six, and W_{RR}, W_{GP}, ... , W_{RE} are independent Wiener process.

Thus, P_T can be solved by:

$$
P_T^2 = P_t e^{-\lambda^2(T-t)} + e^{-\lambda^2 T} \int_t^T e^{\lambda^2 s} \theta_P^2(s) ds + \sigma_P^2 \int_t^T e^{-\lambda^2(t-s)} dW_P. \tag{23}
$$

3.3 SIRDP Pricing Model

For the SIRDP model, we exclude the influence of mobility, then we can deduce a SDE of oil prices from equation (8):

$$
dP_t = (\theta_P^3(t) - \lambda^3 P_t)dt + \sigma_P^3 dW_P \tag{24}
$$

where

$$
\theta_P^3(t) = k_0 + k_I I_t + k_R R_t + k_D D_t \tag{25}
$$

and

$$
\begin{aligned}
S_t &= S_{t_0} - \beta I_{t_0} S_{t_0} (t - t_0) \\
D_t &= D_{t_0} + \delta I_{t_0} (t - t_0) \\
R_t &= R_{t_0} + \gamma I_{t_0} (t - t_0) \\
I_t &= 1 - S_t - R_t - D_t.
\end{aligned} \tag{26}
$$

In this case, the spot prices of crude oil P_T can be solved as:

$$
P_T = P_t e^{-\lambda^3(T-t)} + e^{-\lambda^3 T} \int_t^T e^{\lambda^3 s} \theta_P^3(s) ds + \sigma_P^3 \int_t^T e^{-\lambda^3(t-s)} dW_P. \tag{27}
$$

3.4 Future Pricing

We could get the future prices of the crude oil by taking the expectation of the spot prices at maturity T:

$$F(P_t, T) = \mathbb{E}_t(P_T)$$

For the SIRDMP pricing model with single composite mobility, future price can be shown to be:

$$\begin{aligned} F_1(P_t, \alpha_t, I_t, t, T) &= e^{-\lambda^1(T-t)}P_t + (1 - e^{-\lambda^1(T-t)})c_1 + (e^{-b(T-t)} - e^{-\lambda^1(T-t)})\frac{K_\alpha(\alpha_t - \theta_\alpha)}{\lambda^1 - b} \\ &+ e^{-\lambda^1 T}K_I I_t \int_t^T e^{\lambda^1 s} f(s)ds + e^{-\lambda^1 T}(K_D\delta + K_R\gamma)I_t \int_t^T \int_t^s e^{\lambda^1 s} f(\tau)d\tau ds \end{aligned} \tag{28}$$

where the above functions and constants have the form:

$$\begin{aligned} c_1 &= \frac{K_0 + K_D D_t + K_R R_t + K_\alpha \theta_\alpha}{\lambda^1}, \\ f(s) &= \exp\big((a_0 - \delta - r + \theta_\alpha \hat{a}_1 h)(s-t) + \frac{1}{2}\sigma_\epsilon^2(s-t)^2 - \frac{\hat{a}_1}{b^2}(\alpha_t - \theta_\alpha)(1 - e^{bh})(1 - e^{-b(s-t)}) \\ &+ \frac{\hat{a}_1^2 h^2}{b^2}((s-t) + \frac{1}{2b}(1 - e^{-2b(s-t)}) - \frac{2}{b}(1 - e^{-b(s-t)}))\big). \end{aligned}$$

For the SIRDMP pricing model with multiple mobility, future price can be shown to be:

$$\begin{aligned} F_2(P_t, \alpha_t^M, I_t, t, T) &= e^{-\lambda^2(T-t)}P_t + (1 - e^{-\lambda^2(T-t)})c_2 + \sum_{i=1}^{6}(e^{-b^i(T-t)} - e^{-\lambda^2(T-t)})\frac{K_\alpha^i(\alpha_t^i - \theta_\alpha^i)}{\lambda^2 - b^i} \\ &+ e^{-\lambda^2 T}K_I I_t \int_t^T e^{\lambda^2 s} f(s)ds + e^{-\lambda^2 T}(K_D\delta + K_R\gamma)I_t \int_t^T \int_t^s e^{\lambda^2 s} f(\tau)d\tau ds \end{aligned} \tag{29}$$

where the above functions and constants have the form:

$$\begin{aligned} c_2 &= \frac{\hat{K}_0 + \hat{K}_D D_t + \hat{K}_R R_t + \hat{K}_\alpha^T \theta_\alpha^M}{\lambda^2}, \\ f(s) &= \sum_{i=1}^{6} \exp\big((a_0' - \delta - r + \theta_\alpha^i \hat{a}_1^i h)(s-t) + \frac{1}{2}(\sigma_\epsilon')^2(s-t)^2 - \frac{\hat{a}_1^i}{{b^i}^2}(\alpha_t^i - \theta_\alpha^i)(1 - e^{b^i h})(1 - e^{-b^i(s-t)}) \\ &+ \frac{(\hat{a}_1^i)^2 h^2}{{b^i}^2}((s-t) + \frac{1}{2b^i}(1 - e^{-2b^i(s-t)}) - \frac{2}{b^i}(1 - e^{-b^i(s-t)}))\big). \end{aligned}$$

where b^i is the i^{th} term from $\hat{b} = \begin{bmatrix} b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \end{bmatrix}^T$, θ_α^i is the i^{th} term from $\theta_\alpha^M = \begin{bmatrix} \theta_{RR} & ... & \theta_{RE} \end{bmatrix}^T$, K_α^i, $\hat{a}_1^i$, α_t^i, are the i^{th} term from $\hat{K}_\alpha^T$, $\hat{a}_1^T$,α_t^M, θ_α^M, respectively.

For the SIRDP pricing model, future price is shown to be:

$$
\begin{aligned}
F_3(P_t, I_t, R_t, D_t, S_t, t, T) = e^{-\lambda^3(T-t)}P_t + (1 - e^{-\lambda^3(T-t)})(c_3 - c_4 t) \\
+ c_4\big((T - 1/\lambda^3) - (t - 1/\lambda^3)e^{-\lambda^3(T-t)}\big)
\end{aligned}
\tag{30}
$$

where the above constants have the form:

$$
\begin{aligned}
c_3 &= k_0 + k_D D_t + k_R R_t + k_I I_t, \\
c_4 &= I_t\big(k_I(\beta S_t - \delta - \gamma) + k_D\delta + k_R\gamma\big).
\end{aligned}
$$

4 The Performance of the Two Models in Futures Pricing

4.1 Estimation of the Parameters

In order to apply the SIRDMP model and SIRDP model in derivatives pricing, we require to estimate all the parameters in the joint stochastic process (10,11), (18, 19) and process (24). For SIRDMP model with composite mobility, K_0, K_P, K_α, K_I, K_R, K_D are regression coefficients for linear regression (7), and a_0, a_1, σ_ϵ are the coefficients and standard deviation of residuals for equation (3), the values of the parameters are already presented in table 1 and table 3. For the model with multiple mobility, coefficients are exhibited in table 2 and 5. For SIRDMP model, these parameters are estimated by the same way, and the values of k_0, k_P, k_I, k_R, k_D have been shown in table 4.

The mortality rate δ and recovery rate γ are the coefficients of the discretized approximations of equation (1):

$$D_t - D_{t-\Delta t} = \delta I_{t-\Delta t} + \eta'_t \tag{31}$$

$$R_t - R_{t-\Delta t} = \gamma I_{t-\Delta t} + \eta''_t \tag{32}$$

And the rate of infection β is the coefficient of regressing B_t on S_t:

$$B_t = \beta S_t + \eta'''_t \tag{33}$$

where η_t, η'_t, η''_t are the error terms.

Furthermore, since the mobility α_t are postulated to follow OU process, we estimate θ_α by directly taking average of mobility over the period for futures pricing, and estimate b by using the linear discretized approximation of SDE (10):

$$\alpha_t - \alpha_{t-\Delta t} = \phi + b(\theta_\alpha - \alpha_{t-\Delta t}) + \eta_t \tag{34}$$

where η_t represents residuals.

Since the stochastic dynamics of oil spot price P_t depends on not only the spot price itself, but also COVID-19 statistics, storage change and mobility, λ is adjusted to fit the future prices.

4.2 Comparison with Other Classic Models

After the crude oil price touching a negative value, the CME and ICE changed Black-Schole model to Bachelier model to price oil futures options [4,5], hence we use the standard Bachelier model for comparison. The oil prices obtain the following stochastic process under Bachelier model:

$$dP_t = \kappa_1 P_t dt + \sigma_1 dW_B \tag{35}$$

Then we can get the solution of the above SDE (35) :

$$P_T = P_t e^{\kappa_1 (T-t)} + \sigma_1 \int_t^T e^{\kappa_1 (T-s)} dW_B \tag{36}$$

The future price is calculated by:

$$F_3(P_t, t, T) = P_t e^{\kappa_1 (T-t)} \tag{37}$$

And κ_1 can be estimated by the regression model:

$$P_t - P_{t-\Delta t} = \zeta + \kappa_1 P_{t-\Delta t} + \hat{\eta}_t \tag{38}$$

where $\hat{\eta}_t$ is the error term.

As shown in equation (36), negative value of oil price can be derived if the standard error σ_1 is large enough. However, as we can see from (37), the sign of the future price is always consistent with the spot price. The standard Bachelier model will always fail for the prediction when oil price is negative but future price is positive as what happened in 2020 April.

In our pricing framework, the stochastic process of the oil price has similar format with mean-reverting process, while the mean part is replaced by a function with information of previous oil prices, COVID-19 statistics and mobility. As a consequence, we also use mean reversion model for comparison. [13] suggests that oil price series displays mean reversion by using long memory techniques, which accords with the choice of mean reversion model. The mean-reverting model process is:

$$dP_t = \kappa_2(\mu - P_t)dt + \sigma_2 dW_M \tag{39}$$

Then we can get an explicit solution of the above equation:

$$P_T = P_t e^{-\kappa_2(T-t)} + \mu(1 - e^{-\kappa_2(T-t)}) + \sigma_2 \int_t^T e^{-\kappa_2(T-s)} dW_M \tag{40}$$

The future price is calculated by:

$$F_4(P_t, t, T) = P_t e^{-\kappa_2(T-t)} + \mu(1 - e^{-\kappa_2(T-t)}) \tag{41}$$

The parameter μ can be estimated by the average of oil prices, and speed of mean reversion κ_2 can be estimated by the regression model:

$$P_t - P_{t-\Delta t} = \psi + \kappa_2(\theta_P - P_{t-\Delta t}) + \tilde{\eta}_t \tag{42}$$

where $\tilde{\eta}_t$ is the error term.

As shown in (40) and (41), mean-reverting model supports negative spot price and inconsistency of the sign between spot price and future price.

Table 6: Estimation of parameters for SIRDMP model, SIRDP model, Bachelier model and mean-reverting model.

δ	γ	β	λ^1	λ^2	λ^3
8.9979e-05	0.0068	0.0678	0.20	0.18	0.05
θ_α	b	κ_1	κ_2	μ	h
-0.2648	0.0141	-0.0648	0.0648	40.9281	0.0137
θ_{RR}	θ_{GP}	θ_{PA}	θ_{TS}	θ_{WP}	θ_{RE}
-0.1937	-0.0592	0.1319	-0.3260	-0.2937	0.0959
b_1	b_2	b_3	b_4	b_5	b_6
0.0406	0.1000	0.0110	0.0332	0.1056	0.0753

Table 6 summarize the estimation of parameters of SIRDMP model, SIRDP model, Bachelier model and mean reversion model. Parameters displayed in table 1-5 have not been included in

Figure 2: Performance of SIRDMP model, SIRDP model, Bachelier model and mean-reverting model in pricing futures with maturity 2021-04-21, 2021-08-21, 2021-12-21, respectively.

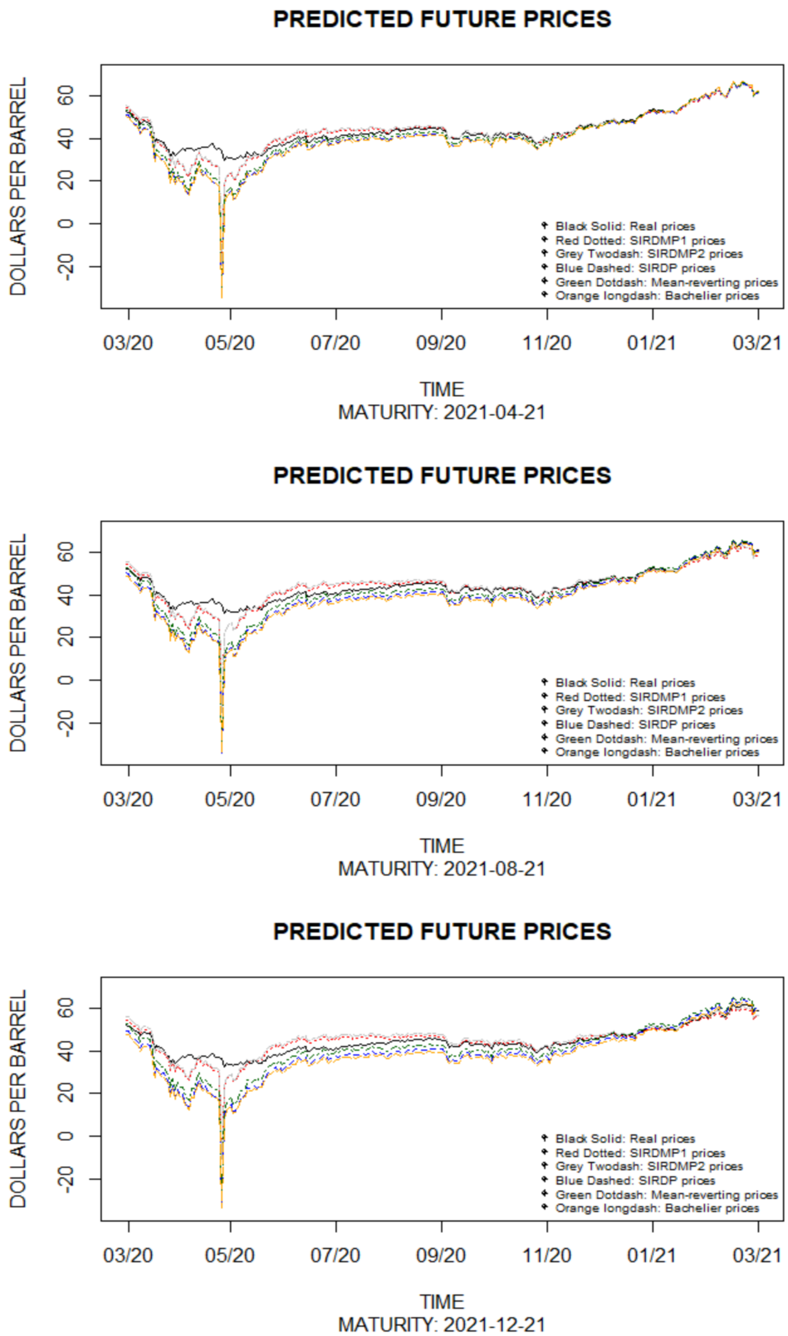

Table 7: Summary statistics on the four models' pricing errors in predicting future prices.

Model	SIRDMCP^1	SIRDMCP^2	SIRDCP	Bachelier	Mean-Reverting
			Maturity: 2020-04-21		
MSE	-0.8228	-0.9960	-4.0998	-4.2244	-3.0252
RMSE	4.3251	4.4938	7.2792	7.4503	6.4502
			Maturity: 2020-08-21		
MSE	-0.5273	-0.7509	-4.8963	-5.7768	-3.3072
RMSE	4.1066	4.2898	8.0233	8.5447	6.7009
			Maturity: 2020-12-21		
MSE	-0.0106	-0.2781	-5.4259	-6.6229	-3.3288
RMSE	3.9373	4.1122	8.7491	9.4445	7.0574

MPE is the mean pricing error in dollars: $MPE = \frac{1}{N}\sum_{n-1}^{N}(\hat{F_n} - F_n)$
RMSE is the root mean squared error in dollars: $RMSE = \sqrt{\frac{1}{N}\sum_{n-1}^{N}(\hat{F_n} - F_n)^2}$
SIRDMP^1 is the model with single mobility and SIRDMP^2 is the model with multiple mobility indices.

table 6. Table 7 shows the summary statistics on pricing errors of SIRDMP model, SIRDP model, Bachelier model and mean reversion model in predicting future prices with distinct maturities. SIRDMP^1 is the model with single mobility and SIRDMP^2 is the model with multiple mobility indices. Figure 2 gives visual exhibitions. As shown in table 7, the pricing performances of SIRDMP model are much satisfactory than SIRDP model, Bachelier model and mean reversion model, which shows the contribution of mobility in predicting the oil future prices. And the goodness of prediction is almost same using single mobility or multiple mobility indices in pricing. The magnitudes of mispricing for futures with maturity in 2021-12-21 are only 0.02% for SIRDMP^1 model, while the magnitudes increase to 16.17% and 8.14% for Bachelier model and mean reversion model. The errors of pricing for the two classic models increase as the maturity of the futures contracts is lengthened, while the pricing performances of SIRDMP model is still robust. According to the results reported in this section, the integration of COVID-19 statistics and mobility to oil pricing framework can significantly raise the performance in valuing futures contracts. Our suggested models have high explanatory for oil spot prices and high forecast power for oil futures. The goodness of our proposed model in pricing futures with long maturities also exhibits that the COVID-19 pandemic has long-lasting effect for the oil market.

5 Conclusion

This paper suggests a new framework of combining stochastic SIRD model with crude oil pricing model during the COVID-19 pandemic. We find that the COVID-19 statistics and public mobility significantly enhance the fittings of oil pricing. We model B_t (βS_t), the infection rate of susceptible population, by two ways, one assumes B_t is a function of public mobility, the other simply assumes β is a constant. After that, we raise SIRDMP model and SIRDP model, and get the solution of oil future prices with the proposed models. Compared to other classic models, the pricing performances of SIRDMP model is much more satisfying during the pandemic crisis. The propose that integrating COVID-19 statistics into crude oil pricing is useful for investors to further improve their crude oil trading strategies. We also discover that the COVID-19 statistics are significantly related to the volatility of crude oil, which encourages us to further study the crude oil option pricing during coronavirus disease.

References

[1] Coronavirus vaccinations. `https://ourworldindata.org/covid-vaccinations` accessed on 2021/06/27.

[2] Covid-19 community mobility reports. `https://www.google.com/covid19/mobility/`.

[3] Worldometer. `https://www.worldometers.info/coronavirus/country/us/`.

[4] CME Group 2020. Worldswitch to bachelier options pricing model - effective april 22, 2020. URL:. `https://www.cmegroup.com/notices/clearing/2020/04/Chadv20-171.html`.

[5] ICE 2020. Negative pricing and option model transition. URL:. `https://www.theice.com/publicdocs/circulars/20057.pdf`.

[6] Claudiu Albulescu. Coronavirus and oil price crash. *Available at SSRN 3553452*, 2020.

[7] Louis Bachelier. Théorie de la spéculation. In *Annales scientifiques de l'École normale supérieure*, volume 17, pages 21–86, 1900.

[8] Norman TJ Bailey et al. *The mathematical theory of infectious diseases and its applications*. Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE., 1975.

[9] Fischer Black. The pricing of commodity contracts. *Journal of financial economics*, 3(1-2):167–179, 1976.

[10] EIA. Cushing, wti spot price fob. `https://www.eia.gov/dnav/pet/hist/rwtcD.htm`.

[11] Jesús Fernández-Villaverde and Charles I Jones. Estimating and simulating a sird model of covid-19 for many countries, states, and cities. Technical report, National Bureau of Economic Research, 2020.

[12] Rajna Gibson and Eduardo S Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance*, 45(3):959–976, 1990.

[13] Luis A Gil-Alana and Manuel Monge. Crude oil prices and covid-19: Persistence of the shock. *Energy Research Letters*, 1(1):13200, 2020.

[14] Giulia Giordano, Franco Blanchini, Raffaele Bruno, Patrizio Colaneri, Alessandro Di Filippo, Angela Di Matteo, and Marta Colaneri. Modelling the covid-19 epidemic and implementation of population-wide interventions in italy. *Nature medicine*, 26(6):855–860, 2020.

[15] Sha He, Sanyi Tang, Libin Rong, et al. A discrete stochastic model of the covid-19 outbreak: Forecast and control. *Math. Biosci. Eng*, 17(4):2792–2804, 2020.

[16] Bo Huang, Jionghua Wang, Jixuan Cai, Shiqi Yao, Paul Kay Sheung Chan, Tony Hong-Wing Tam, Ying-Yi Hong, Corrine W Ruktanonchai, Alessandra Carioli, Jessica R Floyd, et al. Integrated vaccination and physical distancing interventions to prevent future covid-19 waves in chinese cities. *Nature Human Behaviour*, 5(6):695–705, 2021.

[17] Kenji Karako, Peipei Song, Yu Chen, and Wei Tang. Analysis of covid-19 infection spread in japan based on stochastic transition model. *Bioscience trends*, 2020.

[18] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[19] Lu Liu, En-Ze Wang, and Chien-Chiang Lee. Impact of the covid-19 pandemic on the crude oil and stock markets in the us: A time-varying analysis. *Energy Research Letters*, 1(1):13154, 2020.

[20] Stephen E Moore and Eric Okyere. Controlling the transmission dynamics of covid-19. *arXiv preprint arXiv:2004.00443*, 2020.

[21] Paresh Kumar Narayan. Oil price news and covid-19—is there any connection? *Energy Research Letters*, 1(1):13176, 2020.

[22] Afees A Salisu, Godday U Ebuh, and Nuruddeen Usman. Revisiting oil-stock nexus during covid-19 pandemic: Some preliminary results. *International Review of Economics & Finance*, 69:280–294, 2020.

[23] Eduardo S Schwartz. The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of finance*, 52(3):923–973, 1997.

[24] Chanjuan Sun and Zhiqiang Zhai. The efficacy of social distance and ventilation effectiveness in preventing covid-19 transmission. *Sustainable cities and society*, 62:102390, 2020.