TESTING FOR ENDOGENEITY OF COVID-19 PATIENT ASSIGNMENTS

C. Gourieroux*, A. Djogbenou†, J. Jasiak‡

November 13, 2020

Abstract

A considerable number of individuals infected by COVID-19 died in self-isolation. This paper uses a graphical inference method to examine if patients were endogenously assigned to self-isolation during the early phase of COVID-19 epidemic in Ontario. The endogeneity of patient assignment is evaluated from a dependence measure revealing relationships between patients' characteristics and their location at the time of death. We test for absence of assignment endogeneity in daily samples and study the dynamic of endogeneity. This methodology is applied to patients' characteristics, such as age, gender, location of the diagnosing health unit, presence of symptoms and underlying health conditions.

Keywords: COVID-19, Patient Assignment, Endogenous Selection, Conditional Dependence, Value of Human Life, Audit.

*University of Toronto, Toulouse School of Economics and CREST,*e-mail*: gouriero@ensae.fr

†York University, Canada, *e-mail*: daa@yorku.ca

‡York University, Canada, *e-mail*: jasiakj@yorku.ca (corresponding author)

The authors acknowledge the support of Ontario, Ministry of Health and Long-Term Care. The views expressed in this article are the views of the authors and do not necessarily reflect those of Ontario. This research has been approved by the Human Participants Review Committee of the Office of Research Ethics, York University.

The authors thank A. Melino, P. Tuvaandorj and anonymous referees for helpful comments.

C. Gourieroux acknowledges the financial support of Agence Nationale de la Recherche (ANR-COVID) grant ANR-17-EURE-0010

1 Introduction

In all countries infected by COVID-19, a considerable number of people died while self-isolating at home. This paper investigates if patient assignments to self-isolation instead of medical care might have contributed to the deadly outcome of the infection during the early phase of the epidemic in Ontario.

In Ontario, 1434 deaths due to COVID-19 were reported between January 23 and May 05, 2020. Among these, 1116 deaths occurred after at least one day of (self)isolation and 318 in hospital. Among the individuals deceased in self-isolation, 262 were diagnosed after death (post mortem) and 854 were diagnosed before they died. The total number of individuals diagnosed prior to death and deceased either in isolation or in hospital, is 1172. These counts are based on the records of the Public Health Ontario (PHO) database containing individual records of 18722 patients diagnosed with COVID-9 over the period of the first 104 days of the epidemic between January 23 and May 05, 2020. The database IPHIS-REPORT.csv combines records from the iPHIS (integrated Public Health Information System) and CORES (Toronto Public Health Coronavirus Rapid Entry System).

The deceased patients and/or their health care providers, at some point, made a choice between being self-isolated at home or hospitalized. In the case of elderly patients, the families might have been involved in making this decision. In particular, 262 deceased who were not diagnosed prior to death, might have decided by themselves to stay home. Our records do not contain information on whether they sought medical assistance. However, the remaining deceased individuals came into contact with health care providers at least once when tested positive for COVID-19. Therefore, their choice between self-isolation and hospitalization was, if not determined, at least influenced by the health care providers. Due to the presumed role of health care providers, the location of individuals prior to their death will be referred to as the "patient assignment".

At the beginning of the epidemic, little was known of the effects of COVID-19, thus explaining why many patients remained in self-isolation at home, or were sent back home after seeing a doctor. Some of those patients with apparently mild symptoms experienced sudden deterioration of health or even death over a short period of time. Another potential explanation of why only 22% of the deceased died in hospital is the high cost of medical care and/or a concern about an anticipated shortage of beds in the intensive care units (ICU) during the upcoming peak [see e.g. Robert et al. (2011) for the impact of ICU admission on mortality, Palmer et al. (2014) on the Health System Funding Reform and Appendix 4]. During the early phase of the epidemic and under the threat of shortages of ventilators,

medical staff might have felt pressured to only hospitalize patients with severe symptoms.

The aim of our paper is to give ex-post insights on patient assignments to self-isolation during the early phase of the epidemic in Ontario. We examine the endogeneity of assignments by revealing relationships between the characteristics of deceased patients and the states of self-isolation or hospitalization. We describe a dependence measure that allows us to quantify the endogeneity, and study its evolution over time. This methodology is applied to qualitative or quantitative characteristics of patients, such as age, gender, location of the diagnosing health unit, presence of symptoms and underlying health conditions. We test the null hypothesis of absence of endogeneity, which is equivalent to testing the independence between assignment to self-isolation and the deceased patients' characteristics.

The endogeneity analysis applied to daily data on infections in Ontario over the early phase of COVID-19 provides little evidence for endogeneity of daily patient assignments to isolation. This suggests that during the early phase of the epidemic, the daily decisions on remaining in self-isolation were mostly unrelated to patients' characteristics. Nevertheless, we find evidence of endogeneity of daily assignments to isolation for patients over 80 years old and for asymptomatic patients. Over the entire observational period, the absence of endogeneity is rejected for age over 50, absence of symptoms, and the Ottawa location of diagnosing health unit.

The patient assignment also appears unrelated to the increasing demand for ICU beds. We examine the pressure on ICU beds and show that the point-in-time (PIT) difference of counts of available and occupied beds underestimates the potentially needed number of beds. We adjust this PIT measure by adding patients who remained in self-isolation until their death and account for the exponential increase in the number of infections during the early phase of the epidemic.

The paper is organized as follows. In Section 2, we quantify the endogeneity of patient assignments by introducing a dependence measure and test for the absence of endogeneity. In Section 3, this approach is applied to the data on individuals with various characteristics. In Section 4, we examine the pressure on the ICU system. Section 5 concludes. Appendix 1 describes the measure of dependence and the test of independence based on a contingency table. Appendix 2 gives alternative interpretations of the dependence measure. Additional summary statistics are provided in Appendix 3. Appendix 4 relates our outcomes to the structural model of patient care in health economics.

2 The endogeneity of patient assignment

The assignment Y to self-isolation or medical care is a binary variable, which takes values 1 and 0 for isolation and hospitalization, respectively. The state of death is the conditioning variable, denoted by D, which takes values 1 and 0 for death and life, respectively. The qualitative or quantitative explanatory variables Z are the characteristics of patients, such as age, gender, and location of the diagnosing health unit. If the assignment Y to self-isolation is random, for example, due to a diagnosis error caused by the lack of experience with COVID-19, then variables Y and Z are independent given $D = 1$. Otherwise, the assignment exhibits endogeneity, which is evaluated from the dependence measure discussed below.

2.1 The dependence measure

Let us consider the variable Y indicating where the patient was assigned at the time of death:

$$Y = \begin{cases} 1, & \text{if he/she was hospitalized,} \\ 0, & \text{if he/she was isolated.} \end{cases}$$

Let us introduce a binary characteristic Z of a patient. This can be the gender (male or female), an indicator of age (age less than 70 years or more than 70 years), or the location (Toronto area or the rest of Ontario). Then, from the observations on Y_i, Z_i conditional on $D_i = 1$, $i = 1, ..., n$ at time t (day or week), we compute a sample-based 2×2 contingency table:

$$\hat{P}_t = \begin{pmatrix} \hat{p}_{00t} & \hat{p}_{10t} \\ \\ \hat{p}_{01t} & \hat{p}_{11t} \end{pmatrix}, \tag{2.1}$$

where $\hat{p}_{00t}$ is the proportion of old individuals who died while self-isolated, $\hat{p}_{10t}$ is the proportion of young individuals who died while self-isolated, $\hat{p}_{01t}$ is the proportion of old individuals who died hospitalized, and $\hat{p}_{11t}$ is the proportion of young individuals who died hospitalized, in the population of deceased at time t. Thus, $\hat{P}_t$ provides an approximation of the joint distribution of (Y, Z) given death $D = 1$ at time t.

If a patient assignment is random, the variables Y and Z are independent[1] given death

[1]There is an alternative interpretation of the conditional independence condition. Let us consider a physician who can accurately predict the future severity of the disease denoted by S. Then we have a causal chain : $Z \to S \to Y$, i.e., the individual characteristics have an impact on S, and the assignment decision is based on S. Thus the conditional distribution of Y given S and Z is equal to the conditional distribution of Y given S only, which means that variables Y and Z are independent given $S : Y \perp Z|S$. This is the case when $S = D$ that is discussed in our paper.

$D = 1$. The conditional independence is denoted by $Y \perp Z|D$. Otherwise, there is a conditional dependence that may reveal that patients have been endogeneously assigned to self-isolation. The dependence can change over time, depending on the pressure on a medical care unit.

An empirical measure of (conditional) dependence at time t can be defined from a log-linear probability model applied to the 2×2 contingency tables (see Appendix 1). It is given by:

$$\hat{\gamma}_t = \log\left[\hat{p}_{11t}\hat{p}_{00t}/(\hat{p}_{01t}\hat{p}_{10t})\right], \tag{2.2}$$

as an approximation of its theoretical counterpart :

$$\gamma_t = \log\left[p_{11t}p_{00t}/(p_{01t}p_{10t})\right]. \tag{2.3}$$

This measure separates the concordant observations (1,1), (0,0) and the discordant observations (0,1), (1,0).

The measure of conditional dependence γ_t (also called the measure of causation[2]) is equal to the regression coefficient of variable Z in a dichotomous logit model for Y (and to the regression coefficient of variable Y in a logit model for Z, by symmetry) (see, Appendix A2.1).

The variables Y, Z are conditionally independent at time t, if and only if $\gamma_t = 0$. They are conditionally positively (or negatively) dependent according to the sign of γ_t.[3]

We trace out the daily evolution of $\hat{\gamma}_t$ and the associated confidence intervals [see, Appendix 1 for the derivation of an asymptotically valid confidence interval at level 95%] to evaluate the endogeneity of patient assignments over time. This asymptotically valid confidence interval used in this paper, has a simple analytical expression and is valid under the independence hypothesis and its alternative. We test graphically the null hypothesis

$$H_{0t} : \{\gamma_t = 0\} \quad \text{against} \quad H_{1t} : \{\gamma_t \neq 0\} \tag{2.4}$$

by verifying if the asymptotically valid confidence interval at level 95% contains 0, which is a test for absence of endogeneity at date t.[4] Thus we do not average the conditional dependence measure with respect to the threshold as proposed in Chatterjee (2020).

[2]See e.g. Pearl (2009) for a survey on causal inference and the distinction between measure of association (marginal dependence) and measure of causation (conditional dependence).

[3]For a binary variable Z, the selected indicator of outcome has an effect on γ. If Z for outcomes 1, 0 is changed to 1-Z, for outcomes 0, 1 indicator, then the sign of gamma changes to $-\gamma$.

[4]We check if the condition for asymptotic validity of the test given in A.1.2 is satisfied. Alternatively, the quality of the asymptotic approximation can be assessed by applying a permutation test with simulations performed under the null.

2.2 Age Effect

Let us first consider an age indicator Z_a, with thresholds of age equal to $a = 50, 60, 70, 80$. Variable $Z_a = 1$, if the age is greater than a, and 0, otherwise. Figure 1 presents the daily series $\hat{\gamma}_{t,a}$ for patients above each threshold of age along with the evolution of their confidence intervals.

Table 1 shows the counts of assignments among the 1434 deceased by May 04, 1116 of whom died in self-isolation and 318 in hospitals.

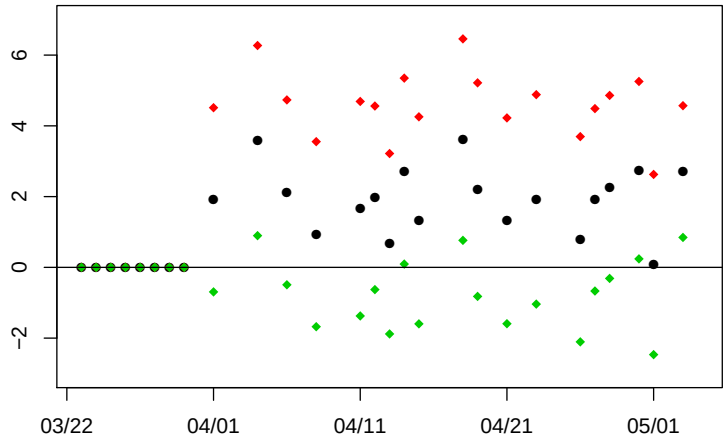
Table 1: Joint Counts: Isolation, Age, Given Death.

assignment	> 50	> 60	> 70	> 80
isolation and $>$ age limit	1080	1043	961	781
isolation and $<$ age limit	36	73	155	335
hospital and $>$ age limit	256	203	138	74
hospital and $<$ age limit	62	115	180	244

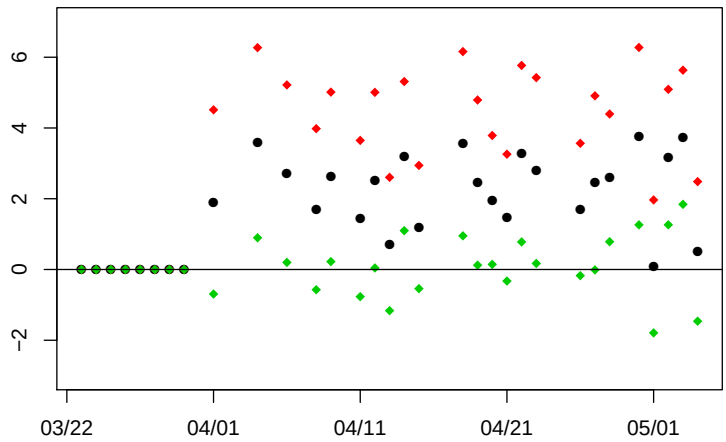
Figures 1 and 2 display the daily dependence measures along with their confidence intervals over the period of March 22 to May 05, 2020. The coefficients are reported on days with a sufficient number of observations, i.e. a sufficiently high daily count of deaths. At the beginning of the epidemic, there are many missing values. For this reason, the figures illustrate the period starting from March 22, 2020.

We observe positive values of conditional dependence.[5] From the interpretation of the sign of dependence measure γ in terms of conditional probabilities (see, Appendix A2.2), it follows that the probability to die in isolation is higher for an older individual than for a younger individual. This effect is observed for all age thresholds with almost equal values of $\hat{\gamma}_t$ on average. It is important to analyze the lower bounds of the dated confidence intervals, which are rather wide due to occasionally limited counts of daily deaths. For $a = 50$, the dated confidence interval contains zero, and the null hypothesis of no endogeneity on day t is not rejected. However, when the threshold of age increases, for example to $a = 80$, at the end of the observational period (end of April until May 05, 2020), the null hypothesis of no endogeneity at day t cannot be rejected.

[5]The condition of asymptotic validity of the Gaussian approximation [see, Appendix A1.2] is satisfied except on days before April 11, on April 26, 29 and May 01.

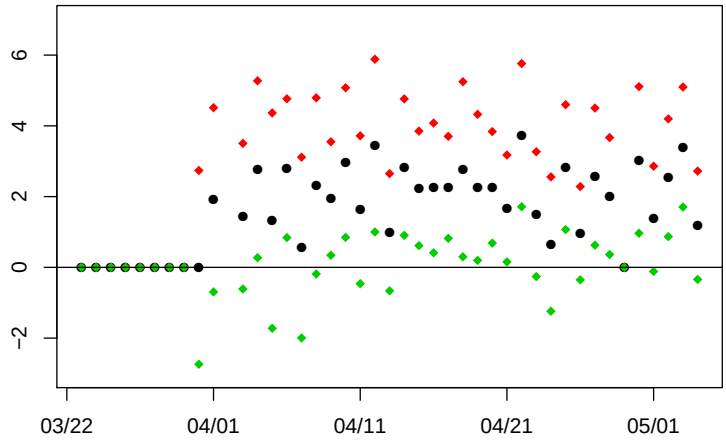


age > 50

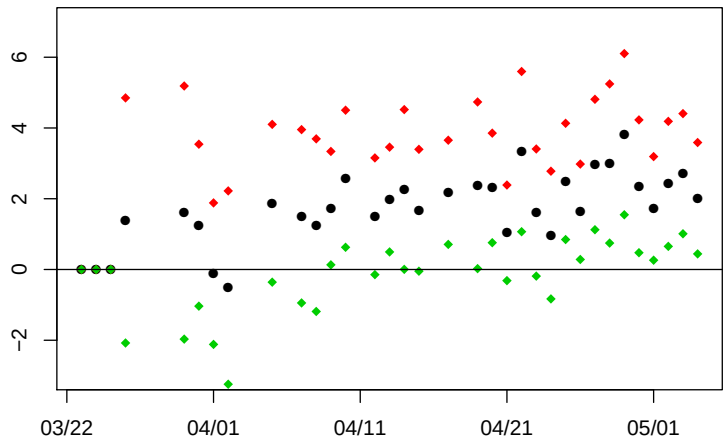


age > 60

Figure 1: Dated Conditional Dependence Measure, Z=Age. The black dots indicate the values of the dependence measure given death $\hat{\gamma}_t$ when the thresholds of age equal to 50 and 60. The diamonds indicate the lower and upper bounds of the confidence intervals. The dots on the line of zero without confidence intervals indicate the missing values of $\hat{\gamma}_t$ due to insufficient numbers of observations and/or zero frequencies.



age > 70



age > 80

Figure 2: Dated Conditional Dependence Measure, Z=Age, Cont. The black dots indicate the values of the dependence measure given death $\hat{\gamma}_t$ when the thresholds of age equal to 70 and 80. The diamonds indicate the lower and upper bounds of the confidence intervals. The dots on the line of zero without confidence intervals indicate the missing values of $\hat{\gamma}_t$ due to insufficient numbers of observations and/or zero frequencies.

The above procedure consists in testing the dated null hypothesis in formula (2.4) date by date. It is a graphical inference method for detecting the occurrence of endogeneity at a given date and selected level 5%, by examining if the lower bound value of the dated confidence interval is above or below 0.[6] The proposed test, as many other routinely performed tests of significance, favours the null hypothesis, i.e. the null hypothesis of no endogeneity at date t and can only detect the presence of sufficiently strong endogeneity at date t.

The slight variation of daily dependence measures over time could be related to the pressure on the ICU [see, Section 4]. As we observe no visible trend in $\hat{\gamma}_t$ for all age thresholds a, even over the period of increased pressure between April 4 and April 10, $\hat{\gamma}_t$ appear independent of the pressure on ICU. This suggests that the standards of medical treatment in the ICU have not changed during the epidemic.

So far, we have considered patients who were diagnosed either before, or after their death. As mentioned in the Introduction, 262 Ontarians died undiagnosed in isolation. The age characteristics of these 262 patients deceased and diagnosed sick of COVID-19 after death are as follows:

quantile	0%	25%	50%	75%	100%
age	24	76	86	92	102

The average age of these patients is 82 years and standard deviation is 14.14. Likely, the elderly deceased were residents of the long term facilities who died of COVID-19 without being diagnosed. They are probably the early victims of the COVID-19 crisis in the retirement homes in Ontario.

Let us now consider only the 1172 patients who were diagnosed before their death (see Table 2). Among these 1172 deceased, 854 individuals died in isolation.

Table 2: Joint Counts of Isolation, Age, Given Death and Diagnosed Before Death

assignment	> 50	> 60	> 70	> 80
isolation and $>$ age limit	830	804	741	611
isolation and $<$ age limit	24	50	113	243
hospital and $>$ age limit	256	203	138	74
hospital and $<$ age limit	62	115	180	244

Figures 12 and 13 displaying the dated dependence measures for the diagnosed individuals are given in Appendix 3. We observe that some dated dependence measures are shifted upwards, but the general conclusions remain unchanged.[7]

[6]The test of the joint hypothesis $\bar{H}_{0,T} : \{\gamma_t = 0,\ \forall t = 1, ..., T\} = \prod_{t=1}^{T} H_{0,t}$ against $\bar{H}_{1,T} : \{\exists t : \gamma_t \neq 0\}$ could be derived from the dated tests at 5%. It will have a different significance level, which can be bounded by the Bonferroni inequality [see Barras et.al. (2010) for methods for avoiding "false discoveries" in a similar, but not equivalent topic]. However, it has no interest in our auditing framework.

[7]The condition of asymptotic validity of the Gaussian approximation [see, Appendix A1.2] is satisfied

2.3 Effect of other individual characteristics and conditioning

The tests of no endogeneity at date t can be applied to other patient characteristics Z. Below, we examine the endogeneity of assignment given the gender, location of diagnosing health unit, presence of symptoms and underlying health conditions.

For each variable Z, we first provide the joint distribution of Y, Z conditional on $D = 1$ over the whole period. Next, we show the plots of the daily measures of dependence. We also consider the effect of changing the conditioning set, when the sub-population of interest includes only the deceased who were diagnosed before death.

2.3.1 Gender Effect

Table 3 below shows the joint counts of deceased in either isolation or hospitalization for variable Z of gender.

Table 3: Joint Counts: Gender, Isolation, Given Death.

assignment	count
isolation and men	483
isolation and women	633
hospital and men	155
hospital and woman	163

Figure 3 below illustrates the conditional dependence between gender and isolation.

We observe that the dated dependence measures are close to zero.[8] Therefore, we can conclude about the absence of endogeneity of patient assignments with respect to gender at the considered dates.

Let us now examine the sample without the 262 patients diagnosed after death presented in Table 4.

Table 4: Joints Counts of Gender, Isolation, Given Death and Diagnosed Before Death.

assignment	count
isolation and men	366
isolation and women	488
hospital and men	155
hospital and woman	163

As shown in Figure 4, the conclusion remains unchanged if the individuals diagnosed post mortem are removed from the sample.[9]

except for days before April 15, on April 17, 18, 23, 26, and between April 29-May 01.

[8]The condition of asymptotic validity of the Gaussian approximation [see, Appendix A1.2] is satisfied except for days before April 11, on April 26, 29 and May 01.

[9]The condition of asymptotic validity of the Gaussian approximation [see, Appendix A1.2] is satisfied except for days before April 15, on April 17, 18, 23, 26, and between April 29-May 01.

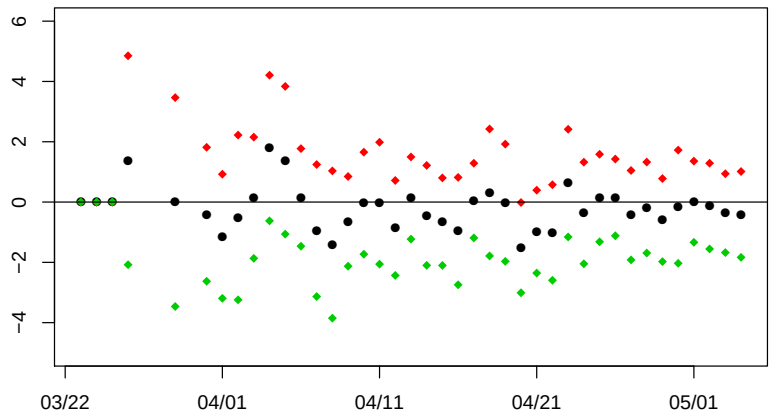


Figure 3: Dated Conditional Dependence Measure, Z=Gender. The black dots indicate the values of the dependence measure conditionally on death $\hat{\gamma}_t$ for gender. The diamonds indicate the lower and upper bounds of the confidence intervals. The dots without confidence intervals on the line of zero indicate the missing values of $\hat{\gamma}_t$ due to insufficient numbers of observations and/or zero frequencies.

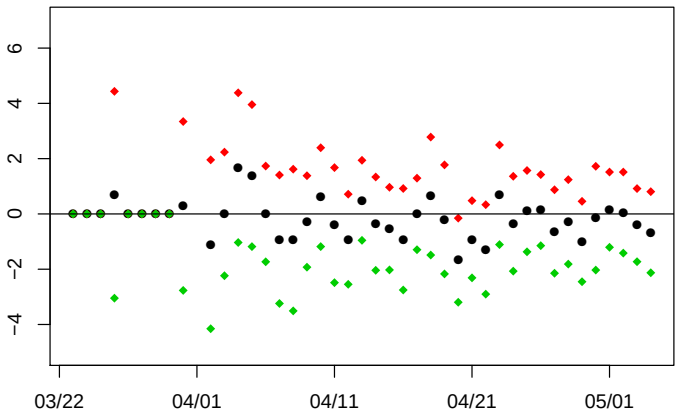


Figure 4: Dated Conditional Dependence Measure, Z=Gender, Diagnosed. The black dots indicate the values of the dependence measure given death and diagnosed before death $\hat{\gamma}_t$ for gender. The diamonds indicate the lower and upper bounds of the confidence intervals. The dots without confidence intervals on the line of zero indicate the missing values of $\hat{\gamma}_t$ due to insufficient numbers of observations and/or zero frequencies.

2.3.2 Area of diagnosis effect

Let us now examine the variable Z of Diagnosis Health Unit Area and the endogeneity tests applied to the regions of Toronto and Ottawa.

More precisely, we consider three binary variables:

$Z_1 = 1$, if the individual is diagnosed in Toronto region, 0, otherwise,

$Z_2 = 1$, if the individual is diagnosed in Ottawa region, 0, otherwise,

$Z = Z_1 + Z_2 = 1$, if the individual is diagnosed either in Toronto, or in Ottawa region.

Table 5: Joint Counts of Location, Isolation, Given Death

assignment	count		count		count
isolation out of Toronto and Ottawa	682	in Toronto	295	in Ottawa	139
isolation in Toronto or Ottawa	434	out of Toronto	821	out of Ottawa	977
hospital out of Toronto and Ottawa	160	in Toronto	147	in Ottawa	11
hospital in Toronto or Ottawa	158	out of Toronto	171	out of Ottawa	307

The counts in Table 5 are crude summaries of the propagation of disease in these two areas. We observe that both the magnitude and patterns are significantly different. From the Public Health Ontario (2020a) report, it follows that the total number of COVID-19 cases before July 4 was 13,372 in Toronto as compared to 2,109 in Ottawa, with populations of 3,120,358 and of 1,054,656, respectively. The peaks of daily number of deaths, were different too: Toronto had the maximum of 43 deaths on April 19 as compared to 12 deaths on May, 1 in Ottawa.[10]

The plots of dated dependence measures for deceased patients who were diagnosed in the Diagnosis Health Unit Area 3895 of Toronto and the Diagnosis Health Unit Area 2251 of Ottawa are given in Figures 5 and 6 below.[11]

The daily dependence measures for deceased patients who were diagnosed in Ottawa [resp. in either Diagnosis Health Unit Area 3895 of Toronto, or 2251 of Ottawa] are given in Figure 7 below (resp. Figure 14, Appendix 3).

Figures 5 and 6 reveal daily negative dependence for the Toronto location and no significant dated dependence for the Ottawa location. The observed negative dependence implies that, all things equal otherwise, the severe cases have a larger probability to be hospitalized in Toronto than outside Toronto. The dated dependence measures are not significant for the Ottawa region. This is largely a consequence of rather low daily death counts. Figure 7 reveals the aggregation issue when both regions are put together, which can be explained from the interpretation of γ in a dichotomous logit model (see, Appendix 2). The dependence measure in Figures 5 and 6, γ_{1t}, γ_{2t}, say, correspond to dated logit models with regression terms $\gamma_1 Z_1$, and $\gamma_2 Z_2$, respectively. In Figure 7, the γ coefficients correspond to the regression term $\gamma Z = \gamma Z_1 + \gamma Z_2$, assuming implicitly that $\gamma_1 = \gamma_2 = \gamma$.

[10]The other areas with high maximum daily counts of deaths are in general close to Toronto and their peak times are similar. These are: Peel area, maximum count 13 on April 30, for a population of 1,605,952, Durham area, maximum count 9 on April 20, for a population of 712,402, York area, maximum count 7 on April 27 for a population of 1, 225,797. Waterloo area had also a high maximum daily count of death of 9 on April 23 for a population of 584,361 [see the interactive map on Public Health Ontario (2020b)].

[11]The condition of asymptotic validity of the Gaussian approximation is satisfied except on days before April 11, on April 26, 29 and May 01.

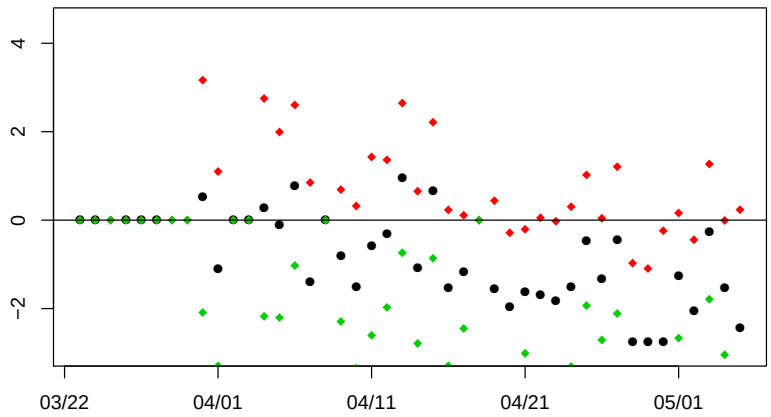


Figure 5: Dated Conditional Dependence Measure, Z=Diagnosed in Toronto. The figure displays the dependence measure given death for deceased patients diagnosed in Toronto. The black dots indicate the values of $\hat{\gamma}_t$, the diamonds indicate the lower and upper bounds of the confidence intervals. The dots without confidence intervals on the line of zero indicate the missing values of $\hat{\gamma}_t$ due to insufficient numbers of observations and/or zero frequencies.

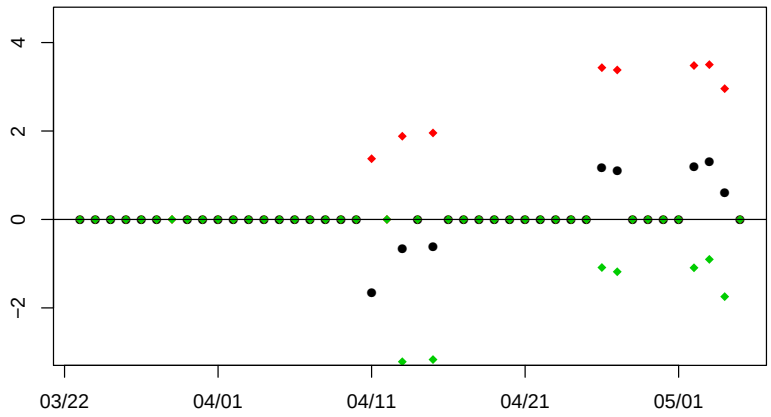


Figure 6: Dated Conditional Dependence Measure, Z=Diagnosed in Ottawa. The figure presents the dependence measure for deceased patients diagnosed in Ottawa. The black dots indicate the values of $\hat{\gamma}_t$, the diamonds indicate the lower and upper bounds of the confidence intervals. The dots without confidence intervals on the line of zero indicate the missing values of $\hat{\gamma}_t$ due to insufficient numbers of observations and/or zero frequencies.

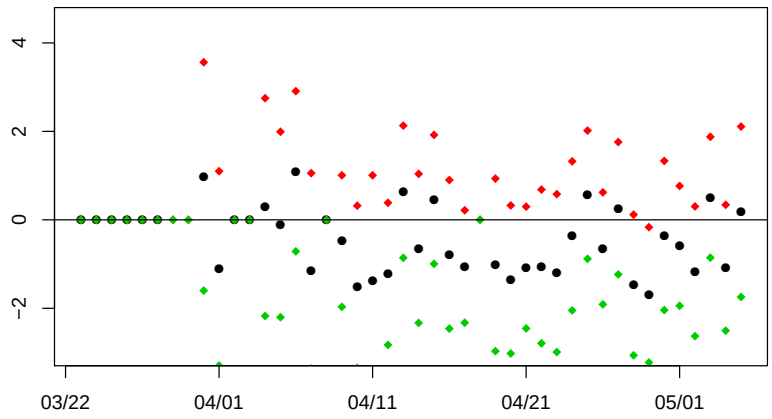


Figure 7: Dated Conditional Dependence Measure, Z=Diagnosis Health Unit Area Toronto and Ottawa. The figures show the daily dependence measure for deceased patients diagnosed in Toronto and Ottawa. The black dots indicate the values of $\hat{\gamma}_t$, the diamonds indicate the lower and upper bounds of the confidence intervals. The dots without confidence intervals on the line of zero indicate the missing values of $\hat{\gamma}_t$ due to insufficient numbers of observations and/or zero frequencies.

Let us verify if the conditional dependence outside of Toronto and Ottawa remains significant in the sample without the 262 deceased patients who were diagnosed post mortem. Table 6 below presents the relevant counts:

Table 6: Joint counts of Location, Isolation, Given Death and Diagnosed Before Death.

assignment	count
isolation, diagnosed out of Toronto and Ottawa	535
isolation, diagnosed in Toronto or Ottawa	319
hospital, diagnosed out of Toronto and Ottawa	160
hospital, diagnosed in Toronto or Ottawa	158

The dated dependence measures are displayed in Figure 14, Appendix 3. We observe that more $\hat{\gamma}_t$ are located above the line of 0, although the confidence interval still contains 0 on all days.

2.3.3 Symptoms Effect

Let us now examine the conditional dependence given the absence of symptoms such as cough, fatigue, fever, headache, shortness of breath, sore throat or other. After removing the missing values, we get a sample of 817 individuals out of whom 470 died in isolation and 347 in hospital. Table 7 below presents the relevant counts:

Table 7: Joint Counts: Symptom, Isolation, Given Death.

assignment	count
isolation, no symptoms	53
isolation symptoms	417
hospital no symptoms	13
hospital symptoms	334

The dated dependence measures are calculated over 10 day intervals starting from March 22 to ensure the validity of results, displayed in Figure 8.

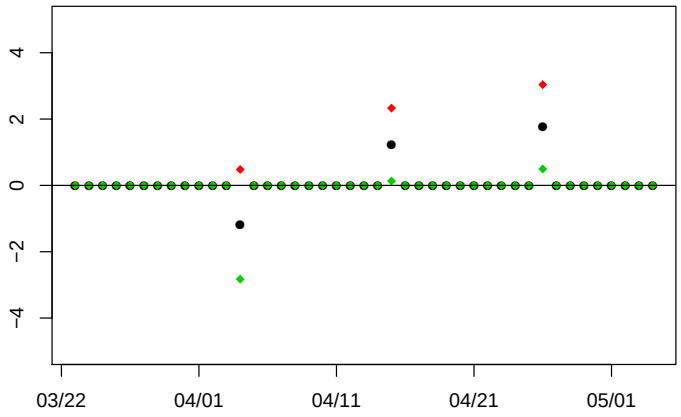


Figure 8: Dated Conditional Dependence Measure, Z=Asymptomatic. The black dots indicate the values of $\hat{\gamma}_t$ in absence of symptoms and conditionally on death. The diamonds indicate the lower and upper bounds of the confidence intervals. The dots without confidence intervals on the line of zero indicate the missing values of $\hat{\gamma}_t$ due to insufficient numbers of observations and/or zero frequencies. Intervals of 10 days are used to have a sufficient number of observations.

We find evidence in favour of endogenous assignment to isolation for the asymptomatic individuals, starting from mid-April.

2.3.4 Effect of underlying health conditions

A similar analysis can be performed with respect to the presence of at least one underlying health condition, such as anemia or hemoglobinopathy, asthma, cancer, cardiovascular condition, UMC, chronic liver disease, COPD, diabetes, immune system disorder, neurological disorder, obesity, postpartum, pregnancy, renal conditions, tuberculosis and other illness. There are 538 deceased individuals to be examined after eliminating the missing values, 277 of whom died isolated and 261 hospitalized (see Table 8).

Table 8: Joint Counts: Health Condition, Isolation, Given Death.

assignment	count
isolation, underlying	14
isolation, no underlying	263
hospital, underlying	26
hospital, no underlying	235

The dated dependence measures calculated over 10 day intervals to ensure the validity of results, are shown in Figure 9.

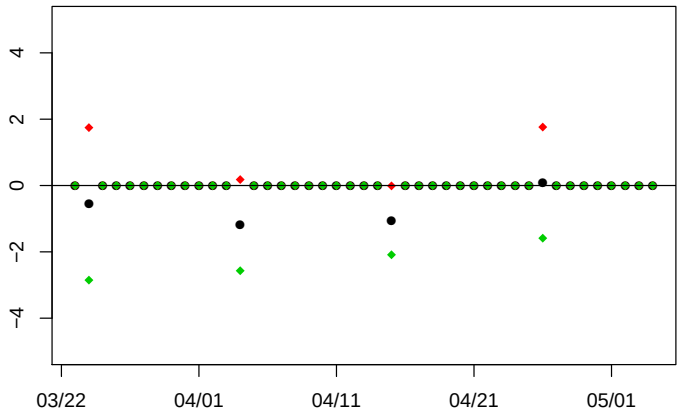


Figure 9: Dated Conditional Dependence Measure, Z=Underlying Health Conditions. The black dots indicate the values of $\hat{\gamma}_t$ in presence of underlying health conditions and given death. The diamonds indicate the lower and upper bounds of the confidence intervals. The dots without confidence intervals on the line of zero indicate the missing values of $\hat{\gamma}_t$ due to insufficient numbers of observations and/or zero frequencies. Intervals of 10 days are used to have a sufficient number of observations.

We find evidence suggesting that in early April (that is the period of pressure - see Section 4 below) patients with underlying health conditions were endogeneously not assigned to isolation.

2.3.5 Ranking of assignment variables

So far, the test was applied to separate variables in order to detect any occurrence of endogenous treatment at date t and to indicate the effects of patient characteristics. A similar approach can be applied to the entire period in order to rank the variables with respect to their explanatory power for patient assignments. We consider a time independent dependence measure γ and test:

$$\bar{H}_0 : \{\gamma_1 = \cdots = \gamma_T = \gamma = 0\} \text{ against } \bar{H}_T : \{\gamma_1 = \cdots = \gamma_T = \gamma \neq 0\}$$

This global test differs from the test of $\bar{H}_{0,T} : \{\gamma_1 = \cdots = \gamma_T = 0\}$ against $\bar{H}_{1T} = \{$ at least one $\gamma_t \neq 0\}$. In the latter approach, the alternative hypothesis does not assume ex ante the equality of the γ_t measures.

The sample dependence measure computed over the whole period T under the assumption of $\gamma_t = \gamma$ constant over time is denoted by $\hat{\gamma}(T)$ and reported in Table 9 summarizing the effects of various characteristics Z. These values of $\hat{\gamma}(T)$ are all statistically significant at 5%, except for the last variable. The statistical significance is related to the larger number of observations from which these dependence measures are computed, as compared to the daily dependence measures.

Table 9: Classification of Variables by $|\hat{\gamma}(T)|$ over the Whole Period.

Z	$\hat{\gamma}(T)$	std error
Age $>$ 60	2.209	0.186
Age $>$ 70	2.146	0.151
Age $>$ 50	2.125	0.250
Age $>$ 80	2.115	0.152
Ottawa	1.369	0.319
Symptoms	1.183	0.318
Toronto	-0.873	0.131
Health condition	- 0.732	0.343
Ottawa-Toronto	0.439	0.132
Gender	*-0.221*	*0.127*

The negative sign of the global dependence measure suggests that patients who were either diagnosed in Toronto or had some underlying health conditions, were endogenously assigned to hospitalization rather than to isolation.

The above ordering by conditional dependence is based on the analysis with a single explanatory variable. In the future, it would be insightful to examine the dependence with several characteristics considered jointly and include the cross-effects of gender and age, for example. This would lead to a new class of cross-variable selection algorithm based on our measure of conditional dependence [see Azadkia, Chatterjee (2020), Section 5 for a similar approach with quantitative regressors and the interpretation in terms of logit regressions in Appendix 2]. At the beginning of the outbreak, due to the limited number of available data (i.e. the limited number of individuals deceased in Ontario) and the missing individual characteristics in the dataset, the validity of a multivariate analysis with crossed Z variables cannot be ensured.[12]

3 Measuring the pressure on ICU beds

Let us now examine ex-post how the numbers of occupied ICU beds would have changed if the patients deceased in isolation were hospitalized and treated for COVID-19. A simple approach to measuring the pressure due to COVID-19 on the health system consists in comparing the

[12]This also explains why machine learning techniques are inapplicable in this context.

number of occupied ICU beds (including beds with ventilators and/or intubation) with the total number of available ICU beds. We adjust this simple measure as follows:

i) we add patients deceased in isolation who might not have been correctly diagnosed and hospitalized early enough;

ii) we account for the exponential increase of infections during the epidemic. This effect is amplified by the long duration (about 14-20 days) of treatment for COVID-19 in the ICU.

The measure of pressure that accounts for the two aspects given above is an ex-post measure of pressure rather than a projection of the demand for ICU beds [see e.g. Gibney (2020), Shoukat et al. (2020)].

i) Adjusting for deceased in self-isolation

For each individual deceased from COVID-19 in self-isolation, the time that patient would have stayed in ICU if he/she were hospitalized, is set equal to 14 days [13]. Then, by summing up those individuals, we find the count of ICU beds that would have been potentially occupied in the past. These ex-post daily counts are plotted in Figure 10 over the period from January 23 to May 01, 2020.

The dashed curve represents the counts of occupied beds and the dotted curve shows the counts of beds potentially occupied if individuals who died in isolation were hospitalized. The solid line represent the sum of both curves.

ii) Adjusting for the dynamics

To adjust for the propagation of the disease, we first consider the daily increase of the count of occupied beds (see Figure 10). This increase is approximately $x\%$ with $x = 0.2$ over the period. We predict the potential total demand (PTD) at horizon of 20 days (3 weeks) by computing the quantity :

PTD at 20 days
=[Count of occupied beds+ Count of potentially occupied beds] x $[1 + 0.2\%]^{20}$.

The series of PTD at 20 days is plotted in Figure 11. By comparing it with the count of occupied ICU beds (dashed line), we get insights on the magnitude of underestimation under the simple approach.

The above measure of PTD still underestimates the true PTD, as it does not account for the infected and undetected individuals [see Gourieroux, Jasiak (2020)] and the fact that an increase of detection effort, i.e. the increase of number of tests performed, may also reveal an additional demand for medical care and beds.

[13] This scenario does not account for the uncertainty in time spent in the ICU. Under such uncertainty, random rationing of beds could occur.

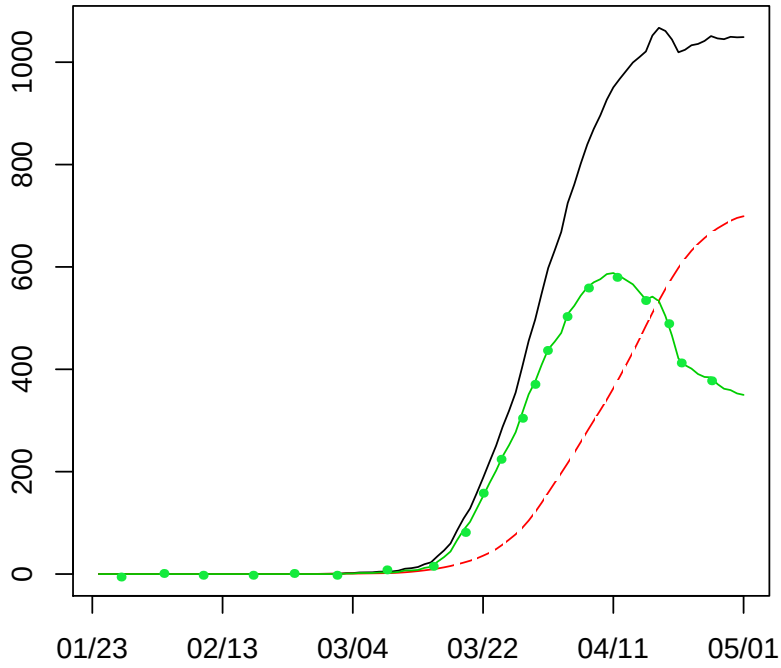


Figure 10: Count of Occupied and Potentially Occupied Beds
dashed line: beds occupied, dotted line: beds potentially occupied, solid line: sum of curves

The daily PTD series have to be compared with the total number of available ICU beds to see if there was a shortage of beds on some days. Indeed "critical care triage to allow for the rationing of scarce ICU resources might be needed" [Shoukat et al. (2020)]. The numbers of critical care beds equipped with ventilators and available for COVID-19 patients was about 2,811 by April 30 [Office of the Premier (2020)]. However, there were only 415 of them on April 4. Next, their number was increased to "reach the level of 1300 before the peak" [CBS News (2020)]. Thus, the potential total supply of ICU beds exceeded the potential demand in Ontario, although the low initial supply of beds created pressure between April 4 and April 10. Under the scenario considered above and based on the data available ex-post, it seems that there were sufficiently many beds available to treat the patients who died in self-isolation during the first wave of the pandemic. However, during that early phase, fewer COVID-19 tests were performed daily, and contact tracing was more limited than later on in Summer 2020. Moreover, some data used for calibration were based on public announcements and could contain error. No data on total daily death counts for inference on undetected infection rates were available either [see Gourieroux, Jasiak (2020)].

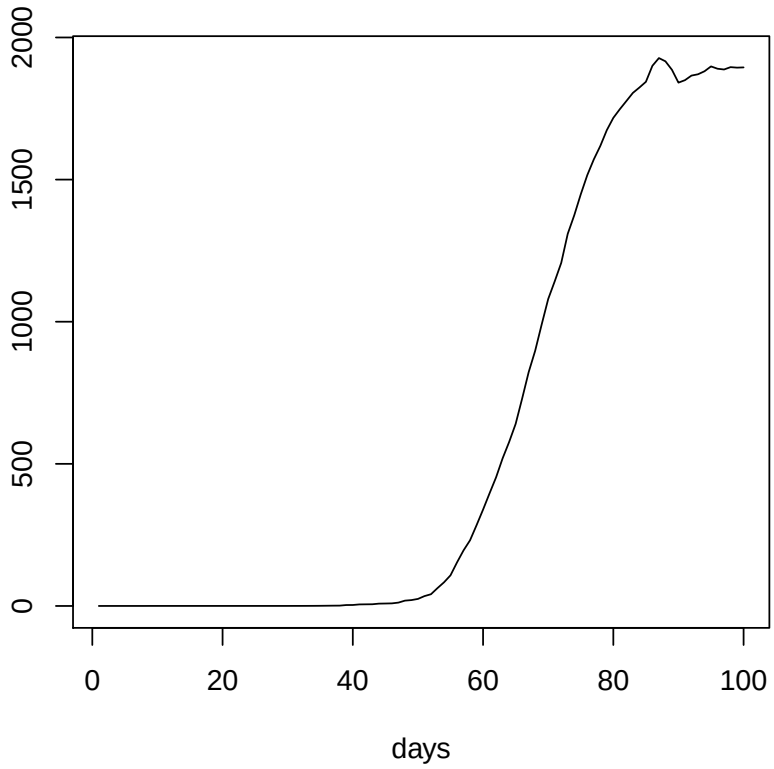


Figure 11: Potential Total Demand at 20 Days

4 Concluding Remarks

The dated measure of conditional dependence allows us to test if patients' characteristics played a role in assigning them to isolation at home or hospitalization at the time of death.

The endogeneity analysis applied to daily individual data on infections in Ontario over the early phase of COVID-19, provides little evidence for endogeneity of daily patient assignments to isolation and suggests that during the early phase of the epidemic, the decisions of remaining in self-isolation were unrelated to patients' characteristics. Nevertheless, after mid-April, we find daily evidence of assignment endogeneity for patients over 80 years old and asymptomatic patients. Over the entire observational period, the absence of assignment endogeneity is rejected for age over 50, the location of diagnosing health unit and the absence of symptoms. There is no evidence of assignment endogeneity with respect to gender. Moreover, patients who were diagnosed in Toronto were more likely to be hospitalized than those who were diagnosed in Ottawa.

The approach developed in this paper can be applied to test at date t for endogeneity of other decisions concerning the individual medical histories.

i) The selection of individuals to be tested for COVID-19 needs to be exogenous. Several countries have decided to test mainly symptomatic individuals and/or those in contact with

symptomatic persons (contact tracing), which is endogenous, while Russia, for example, tested daily random samples of population to determine the rate of asymptomatics in the population.

ii) The average time spent in the ICU seems to decrease on average in the world. The dated dependence measure can be used to test if this is due to enhanced knowledge of COVID-19 and improved treatment (exogenous effect) rather than to endogeneous decision of admitting patients with better prospects of recovery.

iii) By changing the conditioning set, the approach may also be used to detect at which stage the assignment decision is taken, for example, by conditioning on the population of deceased who visited the ER or a walk-in-clinic at the beginning of their infection.

The proposed approach can also be considered as the first-step to deriving patient assignment scores based on the estimation of structural logit models. The second step is beyond the scope of our paper [see, Appendix 4].

The test of endogeneity is important as a doctor may be held liable for causing the patient's loss of a chance to be cured if he/she fails to diagnose the future evolution of the disease [see Government of Canada (2006), subsection "Medical Assistance in Dying-Reporting of Deaths"].

The graphical inference method introduced in this paper can also be used for auditing not only the medical practice. The detection of abnormal endogeneity can be relevant, for example, to detect benefit fraud, the use of insider information by investors, or to assess credit granting decisions.

Appendix 1

Measure of Dependence and Independence Test for 2×2 Contingency Tables.

This appendix describes the test of independence between dichotomous variables, which is an alternative to the chi-square test of independence used, for example, in controlled medical experiments [see Mantel, Haenzel (1985) for early reference]. The independence test is used in this paper to test for conditional independence date by date. For ease of exposition, the time index is omitted in this Appendix. From its logit model interpretation, it follows that the approach can be extended to include more variables (regressors). Let us consider a (2×2) contingency table for binary variables Y and Z:

$$P = \begin{pmatrix} p_{00} & p_{01} \\ \\ p_{10} & p_{11} \end{pmatrix},$$

which is a 2×2 matrix with positive elements that sum up to one. Its sample, i.e. its frequency-based counterpart:

$$\hat{P} = \begin{pmatrix} \hat{p}_{00} & \hat{p}_{01} \\ \\ \hat{p}_{10} & \hat{p}_{11} \end{pmatrix},$$

is computed from n independent observations on binary variables Y and Z. The vector of probabilities is denoted by p and given by:

$p \equiv (p_{00}, p_{10}, p_{01}, p_{11})'$.

Its frequency-based counterpart is denoted by $\hat{p}$.

A1.1 Log-linear probability model of $p_{k,l}$

A log-linear probability model of $p_{k,l},\ k, l = 0, 1$ is obtained from an additive representation of the log-probability in terms of the mean, marginal and cross effects as:

$$\log p_{kl} = \mu + \alpha_k + \beta_l + \gamma_{kl},$$

subject to the constraints :

$$\alpha_1 + \alpha_0 = 0, \beta_1 + \beta_0 = 0, \gamma_{1l} + \gamma_{0l} = 0, l = 0, 1, \gamma_{k1} + \gamma_{k0} = 0, k = 0, 1,$$

and the unit mass restriction [see Nerlove, Press (1973), (1986)].

It follows from simple computation that the expression of $\gamma = \gamma_{11}$ is:

$$\gamma = \log \left[p_{11}\ p_{00} / (p_{01}\ p_{10}) \right]. \tag{a.1}$$

Parameter γ is a measure of dependence between binary variables Y and Z. It can take any value in $(-\infty, +\infty)$. In particular, the measure of dependence is equal to zero, i.e. $\gamma = 0$, if and only if :

$$p_{11}\, p_{00} = p_{01}\, p_{10}$$

$$\iff \det \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} = 0$$

$$\iff \text{the columns of } P \text{ are linearly dependent}$$

$$\iff p_{kl} = p_{k.} p_{.l} \ \ \forall k, l, \text{ where } p_{k.} = p_{k1} + p_{k0}, p_{.l} = p_{1l} + p_{0l},$$

which is the condition of independence for variables Y and Z. Therefore, testing for $\gamma = 0$ is equivalent to testing for independence of Y and Z (see, Section A.3 below).
The measure of dependence γ is positive (resp. negative) if and only if the underlying binary variables Y, Z are positively (resp. negatively) correlated. Indeed, we have :

$$\text{cov } (Y, Z) = E(YZ) - EYEZ$$

$$= p_{11} - p_{1.} p_{.1}, \text{ since Y, Z and YZ are Bernoulli variables,}$$

$$= p_{11} - (p_{11} + p_{10})(p_{11} + p_{01})$$

$$= p_{11}(1 - p_{10} - p_{01} - p_{11}) - p_{10}\, p_{01}$$

$$= p_{11}\, p_{00} - p_{10}\, p_{01}.$$

Therefore, $\text{cov}(Y, Z) > 0, \text{ iff } \gamma > 0.$

A1.2 Asymptotic distribution of $\hat{\gamma}$

Let $\hat{\gamma} = \log[\hat{p}_{11}\ \hat{p}_{00}/(\hat{p}_{01}\ \hat{p}_{10})]$ denote the sample counterpart of γ. For large n, the asymptotic distribution of the vector of probabilities $\hat{p}$ is :

$$\sqrt{n}(\hat{p} - p) \sim N(0,\ \text{diag } (p) - pp'), \tag{a.2}$$

where diag p is the diagonal matrix with the components of p as diagonal elements.
The asymptotic distribution of $\hat{\gamma}$ is derived from the above formula by the δ-method. We get :

$$\sqrt{n}(\hat{\gamma} - \gamma)_{\sim} \left[\frac{1}{p_{00}}, -\frac{1}{p_{01}}, -\frac{1}{p_{10}}, \frac{1}{p_{11}}\right] \sqrt{n}(\hat{p} - p).$$

It follows that the asymptotic distribution of $\hat{\gamma}$ is:

$$\sqrt{n}(\hat{\gamma} - \gamma) \sim N\left[0, \frac{1}{p_{00}} + \frac{1}{p_{01}} + \frac{1}{p_{10}} + \frac{1}{p_{11}}\right]. \tag{a.3}$$

An asymptotically valid confidence interval for γ at level 95 % is:

$$\left(\hat{\gamma} \pm \frac{1.96}{\sqrt{n}} \sqrt{\frac{1}{\hat{p}_{00}} + \frac{1}{\hat{p}_{01}} + \frac{1}{\hat{p}_{10}} + \frac{1}{\hat{p}_{11}}}\right). \tag{a.4}$$

The above asymptotic confidence interval is valid under the conditions required for the validity of the Gaussian approximation of a multinomial distribution. These are: n large and $n\hat{p}_{kl}$ larger than 10, say, for any k, l. Otherwise, Poisson approximations have to be used.

A1.3 Test of the independence hypothesis

If the validity condition for the asymptotic Gaussian approximation is satisfied, the null hypothesis of independence of binary variables Y and Z can be tested. Testing the null hypothesis of independence of Y and Z is equivalent to testing $H_0 : \gamma = 0$ against $H_1 : \gamma \neq 0$. The asymptotically valid test of H_0 at level 5% is as follows. We reject H_0 if

$$\left| \frac{\hat{\gamma}}{(1/\sqrt{n})\left(\sqrt{\frac{1}{\hat{p}_{00}} + \frac{1}{\hat{p}_{01}} + \frac{1}{\hat{p}_{10}} + \frac{1}{\hat{p}_{11}}}\right)} \right| > 1.96$$

and do not reject, otherwise. Equivalently, H_0 is not rejected if the asymptotically valid confidence interval given above contains zero, and it is rejected, otherwise.

Appendix 2

Conditional Interpretation of Dependence Measure and its Sign

Appendix 1 presented interpretations of γ from the joint distribution of Y and Z. This appendix provides alternative interpretations in terms of the conditional distribution of Y given Z (and Z given Y).

A2.1 The Dichotomous Logit Model

Let us consider a dichotomous logit model for variable Y with binary variable Z as the explanatory variable. This model is defined by:

$$P[Y = 1|Z] \equiv [1 + exp(\alpha + \gamma Z)]^{-1}.$$

Therefore we have:

$$log\frac{P[Y = 1|Z]}{P[Y = 0|Z]} = \alpha + \gamma Z.$$

It follows that:

$$log\frac{P[Y = 1|Z = 1]}{P[Y = 0|Z = 1]} = log\frac{P[Y = 1, Z = 1]}{P[Y = 0, Z = 1]} = \alpha + \gamma,$$

$$log\frac{P[Y = 1|Z = 0]}{P[Y = 0|Z = 0]} = log\frac{P[Y = 1, Z = 0]}{P[Y = 0, Z = 0]} = \alpha.$$

Then, by taking the difference, we get:

$$\begin{aligned}\gamma &= log\frac{P[Y=1,Z=1]}{P[Y=0,Z=1]} - log\frac{P[Y=1,Z=0]}{P[Y=0,Z=0]} \\ &= log\frac{p_{11}p_{00}}{p_{10}p_{01}}.\end{aligned}$$

This provides the interpretation of dependence measure γ as a regression coefficient in a logit model for Y with a binary explanatory variable Z.

A2.2 Conditional Probabilities

The sign of dependence measure γ is easy to interpret in terms of conditional probabilities. Let $logit\ x = log\frac{x}{1-x}$ denote the logit transformation. We get:

$$\gamma = logitP[Y=1|Z=1] - logitP[Y=1|Z=0].$$

As the logit transformation is an increasing transformation, it follows that:

$$\gamma > 0 \iff P[Y=1|Z=1] > P[Y=1|Z=0].$$

This equivalence is used for interpreting the empirical results in the text. Also, it is equivalent to:

$$P[Y=0|Z=1] < P[Y=0|Z=0]$$

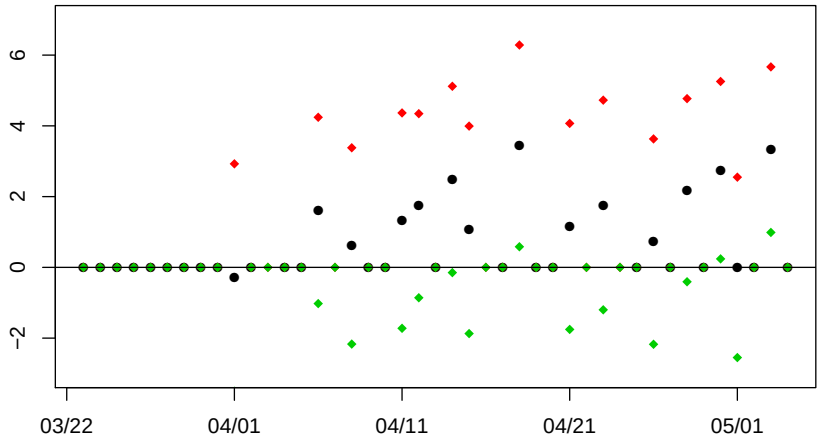
and to

$$P[Y^*=1|Z=1] < P[Y^*=1|Z=0],$$

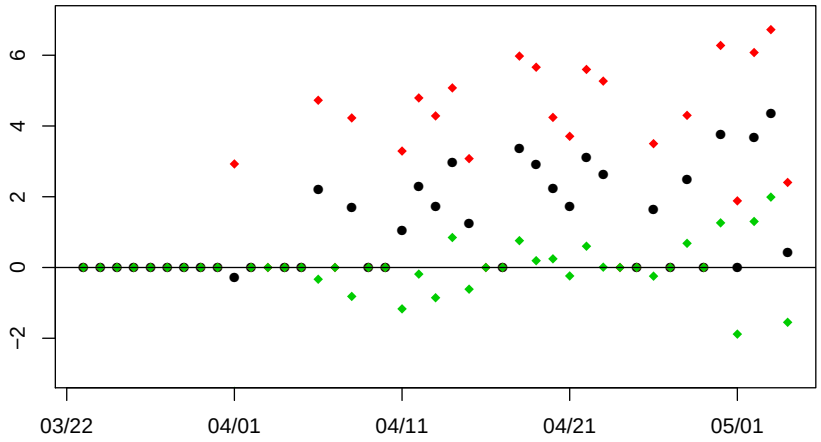
where $Y^* = 1 - Y$,

Appendix 3

Additional Figures

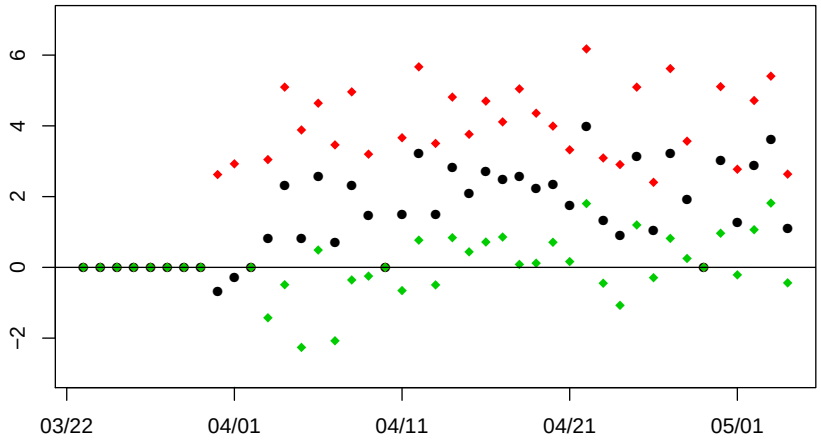


age > 50

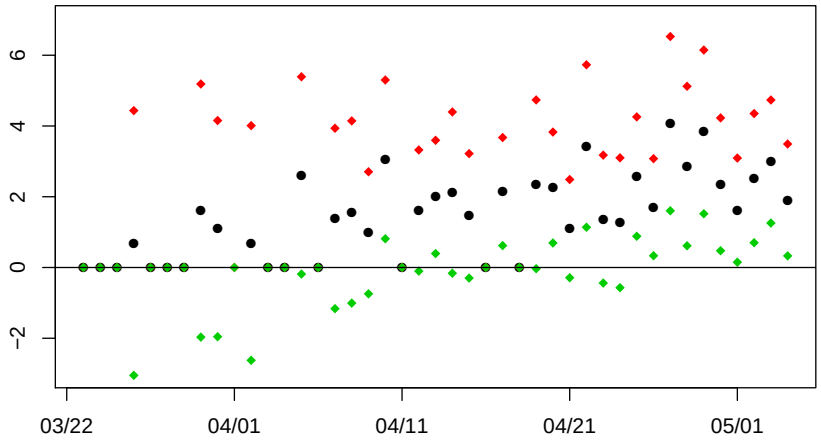


age > 60

Figure 12: Dated Conditional Dependence Measure, Z=Age, Diagnosed. The figures display the daily dependence measure given death and diagnosed before death when thresholds of age equal to 50 and 60. The black dots indicate the values of $\hat{\gamma}_t$, the diamonds indicate the lower and upper bounds of the confidence intervals. The dots without confidence intervals on the line of zero indicate the missing values of $\hat{\gamma}_t$ due to insufficient numbers of observations and/or zero frequencies.



age > 70



age > 80

Figure 13: Dated Conditional Dependence Measure, Z=Age, Diagnosed Cont. The figures display the daily dependence measure given death and diagnosed before death when thresholds of age equal to 70 and 80. The black dots indicate the values of $\hat{\gamma}_t$, the diamonds indicate the lower and upper bounds of the confidence intervals. The dots without confidence intervals on the line of zero indicate the missing values of $\hat{\gamma}_t$ due to insufficient numbers of observations and/or zero frequencies.

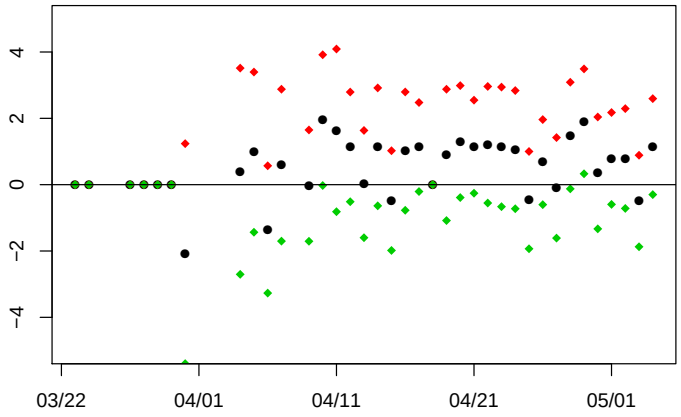


Figure 14: Dated Conditional Dependence Measure, Z=outside of Toronto and Ottawa, Diagnosed. The figures show the daily dependence measure given death and diagnosed before death for patient diagnosed outside of Toronto and Ottawa. The black dots indicate the values of dependence measure $\hat{\gamma}_t$, the diamonds indicate the lower and upper bounds of the confidence intervals. The dots without confidence intervals with contours on the line of zero indicate the missing values of $\hat{\gamma}_t$ due to insufficient numbers of observations and/or zero frequencies.

Appendix 4
Medical and Economic Scores

In this paper, we developed a descriptive tool for scoring the patient assignment decisions, which is a reduced form of a structural model existing in the economic literature. That structural model considers the utility of a surviving patient based on medical and economic arguments.

Under the shortage and rationing of ICU beds, the patient can be selected under either pure medical, or on medical and economic criteria. This approach is used to manage and control health spending based on a cost/effectiveness (i.e. cost benefit) analysis [see e.g. Garber, Phelps (1997), Abelson (2003), Bergstrom (2006), Hall (2007), Zenios (2009)]. We do not discuss the trade-off between the medical and economic aspects, or the ethical rules for validating such techniques [see e.g. Bayles (1970)].

Let us consider patient i with characteristics Z_i. The Health Economics and Ethics literature determine the expected utility for the society of this individual if he/she is admitted to the ICU, and if he/she is not.[14] In a simplified framework, these expected utilities can be decomposed as :

$$U_{1i} = p_1(Z_i)ERL_1(Z_i)\, Val(Z_i), \text{if in } ICU,$$

$$U_{0i} = p_0(Z_i)ERL_0(Z_i)\, Val(Z_i), \text{otherwise,}$$

where $p_k(Z_i)$ is the probability to stay alive in, or out of ICU, $ERL_k(Z_i)$ is the expected residual lifetime when alive, and $Val(Z_i)$ is the economic value per year of the alive patient.

Then the patient can be selected on the basis of its incremental expected utility for the society, that is :

$$\Delta U_i = U_{1i} - U_{0i} = [p_1(Z_i)ERL_1(Z_i) - p_0(Z_i)ERL_0(Z_i)]Val(Z_i).$$

In this expression the term between brackets is purely medical, based on the predictions about the results of the assignment and treatment. The second term $Val(Z_i)$ is economic, known as the price of human life. In this model, it is implicitly assumed for expository purpose and in order to get a simple decomposition of the expected utility that this value per year does not depend on age, whereas the health status is changing with age [see Pliskin et al. (1980)]. This simplifying assumption is usually made in practice. This value is usually estimated indirectly from market data and the willingness-to-pay to stay alive in good health [see e.g. Viscusi, Aldi (2003)].

Then, individual 1 will be selected among the two individuals 1,2, if :

$$\Delta U_1 > \Delta U_2.$$

[14] This utilitarian (or social Darwinian) approach is largely used in public economics for road safety [see e.g. Blaeij et al. (2003), Department of Transportation (2014)], for designing life insurance contracts, or for valuation of damages in the loss-of-chance doctrine law.

The first (medical) component can be estimated from the observations on past treatment effects. For COVID 19, it is known that the age is a significant factor[15] for the survival probability p_1, and also for the ERL, which diminishes with age. The second component is a topic of philosophical and ethical discussions, which are out of the scope of this Appendix [see e.g. Bayles (1978), or Baker et al. (2008) for discussions].

[15]The knowledge of the conditional measure of association provides no information on the measure of association between Y and D.

R E F E R E N C E S

Abelson, P. (2003) : "The Value of Life and Health for Public Policy", The Economic Record, 79, 52-513.

Adler, M. (2020) : "What Should we Spend to Save Lives in a Pandemic? A Critique of the Value of Statistical Life", Covid Economics, 33, 1-45, CEPR Press.

Andersen, T. (1995) : "Adjustment Costs in Price and Quantity Adjustment", Economics Letters, 47, 343-349.

Azadkia, M., and S. Chatterjee (2020) : "A Simple Measure of Conditional Dependence", Arxiv:1910.12327v4.

Baker, R., Chilton, S., Jones-Lee, M., and M., Metcalf (2008) : "Valuing Lives Equallly : Defensible Premise or Unwarranted Compromise ?", Journal of Risk and Uncertainty, 36, 125-138.

Barras, L., Scaillet, O. and R. Wermers (2010) : "False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas", Journal of Finance, 55, 179-216.

Bayles, M. (1997) : "The Price of Life", Ethics, 89, 20-34.

Bergstrom, T. (2006) : "Benefit Cost in a Benevolent Society", American Economic Review, 96, 339-354.

Blaeij de, A., Florax, J., Rietveld, P., and E., Verkoef (2003) : "The Value of Statistical Life in Road Safety : a Meta Analysis", Accident Analysis and Prevention, 35, 973-986.

CBC News (2020) : "Ontario Needs 900 New ICU Beds to Cope with Coming Surge in COVID-19 Patients", April 04.

Chatterjee, S. (2020) : "A New Coefficient of Correlation", Arxiv: 1909.10140v4.

Clarke, S. (2015) : "Medical Malpractice Liability : Canada", Law Library of Congress.

Department of Transportation (2014) : "Guidance on Treatment on the Economic Value of a Statistical Life", June.

Garber, A., and C., Phelps (1997) : "Economic Foundations of Cost-Effectiveness Analysis", Journal of Health Economics, 16, 1-31.

Gibney, N. (2020) : "Critical Care Unit Capacity in Canada : Have we Failed to Advocate Adequatly", Univ. of Alberta, Division of Critical Care Medecine.

Gourieroux, C., and J., Jasiak (2020) : "Time Varying Markov Process with Partially Observed Aggregate Data : An Application to Coronavirus", Journal of Econometrics, forth-

coming.

Government of Canada (2006) : "Health Care System Legislation and Guideline", Section on Guideline, website of the Public Health Agency of Canada.

Government of Canada (2020) : "Canada Health Act", RSC, 1985, c., C-6, Justice Laws Website.

Hall, R. (2007) : "The Value of Life and the Rise of Health Spending", Quarterly Journal of Economics, 39-72.

Keeler, E. (2001) : "The Value of Remaining Lifetime is Close to Estimated Values of Life", Journal of Health Economics, 20, 141-143.

Mangan, M. (1997) : "The Loss of Chance Doctrine : A Small Price to Pay for Human Life", Heinline

Martel, N. and W. Haenszel (1959) : "Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease", J. Natl. Cancer Intst., 22, 719-748.

Ministry of Health Ontario (MOH) (2018) : "Health System Funding Reform, Quality Based Procedures", MOH Website.

Mullbauer, J., and R., Portes (1975) : "Macroeconomic Models with Quantity Rationing", Economic Journal, 85, 788-824.

Nerlove, M., and J., Press (1973) : "Univariate and Multivariate Log-Linear and Logistic Models", Rand Santa-Monica, Prepared for the Economic Development Administration and the National Institutes of Health.

Nerlove, M., and J., Press (1986) : "Multivariate Log-Linear Probability Models in Econometrics", in Mariano, R., ed., Advances in Statistical Analysis and Statistical Computing, 117-171, JAI Press, Greenwich, CT.

Office of the Premier, News (2020) : "Ontario Significantly Expands Hospital Capacity to Prepare for Any COVID-19 Outbreak Scenario", April 16, 1:00pm.

Palmer, K., Agoritsa, T., Martin, D., et al. (2014) : "Activity Based Funding of Hospitals and Impact on Mortality, Readmission, Discharge Destination, Severity of Illness and Volume of Care : A Systematic Review and Meta Analysis", PLOS ONE, 9.

Palmer, K., Brown, A., Evans, J., Marani, H., Russell, K., Martin, D. and N., Ivers (2018) : "Standardising Costs or Standardizing Care ? Qualitative Evaluation of the Implementation and Impact of a Hospital Funding Reform in Ontario, Canada", Health Research Policy and Systems, 16, 74.

Pearl, J. (2009) : "Causal Inference in Statistics : An Overview", Statistics Surveys, 3,

96-146.

Phua, J., et al. (2020) : "Intensive Care Management of Coronavirus Disease 2019 : Challenges and Recommendation", The Lancet, Respiration Medicine, 8, 506-517.

Pliskin, J., Shepard, D., and M., Weinstein (1980) : "Utility Function for Life Years and Health Status", Operations Research, 28, 206-224.

Public Health Ontario (2020a) : "COVID-19 : Epidemiologic Summaries", July. 5.

Public Health Ontario (2020b) : "Ontario COVID-19 Data Tool".

Robert, R., et al. (2011) : "Refusal of Intensive Care Unit Admission Due to a Full Unit Impact on Mortality", American Journal of Respiratory and Critical Care Medicine", 185, issue 10, 1081-1087.

Shoukat, A., et al. (2020) : "Projecting Demand for Critical Care Beds During COVID-19 Outbreaks in Canada", CMAJ Group, May, 192.

Viscusi, W., and J., Aldy (2003) : "The Value of Statistical Life : A Critical Review of Market Estimates Throughout the World", Journal of Risk and Uncertainty, 27, 5-75.

Zenios, S. (2009) : "An Empiric Estimate of the Value of Life : Updating the Renal Dialysis Cost Effectiveness Standard", Value in Health, 12, 80-87.