Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company’s public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Associations between mobility and socio-economic indicators vary across the timeline of the Covid-19 pandemic

Jed A. Long a,*, Chang Ren a, b

a Department of Geography & Environment, Western University, London, Ontario, Canada
b State Key Laboratory of Information Engineering in Surveying Mapping and Remote Sensing, Wuhan University, Wuhan, Hubei, China

1. Introduction

Non-pharmaceutical interventions, such as those imposed by governments worldwide to reduce individual travel and encourage social distancing, were crucial to slowing the spread of Covid-19 in the first wave (Flaxman et al., 2020; Zhang et al., 2020). While regional differences exist in how implemented lockdown restrictions were imposed, in general these have had the effect of reducing international and inter-regional travel (Kang et al., 2020; Wells et al., 2020), widespread increase in home-based working (Reuschke & Felstead, 2020), and reductions in close contacts with others, especially with those outside of one’s household (Ferguson et al., 2020). In the first wave of Covid-19, mobility measures derived from cellular network data (Badr et al., 2020; Chang et al., 2020; Xiong, Hu, Yang, Luo, & Zhang, 2020) and tech company mobility products (Kurita, Sugishita, Sugawara, & Ohkusa, 2021; Noland, 2021; Paez, 2020) were found to be clearly associated with Covid-19 infection rates. However, Covid-19 infection growth rates were not associated with mobility levels outside of the first wave (Gatalo, Tseng, Hamilton, Lin, & Klein, 2020) largely due to mobility levels returning to pre-pandemic levels. During the first wave of Covid-19, changes in baseline mobility levels at the peak of the lockdown were strongly associated with socio-economic indicators in multiple countries (Bonaccorsi et al., 2020; Jay et al., 2020; Lee, Qian, & Schwanen, 2021; Weill, Stigler, Deschenes, & Springborn, 2020) in a way that reflects the luxury nature (Huang et al., 2021) or privilege gap (Dasgupta, Funk, Lazard, White, & Marshall, 2020) of social distancing. As we move into the second (and third)-wave of infections, associations between mobility and socio-economic indicators provide evidence of the ongoing socio-spatial inequities arising from the Covid-19 pandemic (Pareek et al., 2020; Rothwell, 2020; van Dorn, Cooney, & Sabin, 2020) here we study this relationship in the province of Ontario, Canada across the year 2020.

To date studies looking at the associations between aggregate mobility levels and socio-economic variables have been limited in several ways. First, current studies have primarily focused on the first phase of lockdowns occurring during spring 2020, often only looking at a single (or few) time points (Bonaccorsi et al., 2020; Jay et al., 2020; Lee et al., 2021). Given the length of the (ongoing) pandemic, and the dynamic nature of government imposed restrictions aimed to limit spread, it is important to investigate whether these associations have changed over time and in subsequent waves of the pandemic. A second limitation is that present studies use aggregation units that are relatively...
coarse-grained for studying socio-economic relationships. For example, mobile phone data was aggregated to public health regions in England (Lee et al., 2021) and counties in the United States (Dasgupta et al., 2020), and (Bonaccorsi et al., 2020) use a sample of municipalities across Italy based on Facebook data. A final limitation of these studies is that they principally focus their analysis on only a single measure of mobility, which fails to address the fact that human mobility is not easily quantified using a single index (Fillekes, Giannoulis, Kim, Zijlstra, & Weibel, 2019). Thus, it is possible that the observed relationships between socio-economic factors and changes in mobility may differ depending on what measure of mobility is used.

In this study, we explore how three different aggregate mobility measures, captured at a fine geographical scale, are associated with socio-economic factors across 52 weeks of 2020 for Ontario, Canada. We use spatial seemingly unrelated regression (SUR) models (Anselin, 1988) which allows us to account for the spatial structure present in the data and at the same time study how these associations change over time. Our results demonstrate a strong association between changes to baseline mobility and socio-economic conditions during the first wave of the pandemic. We also show that as mobility levels have returned (on aggregate) to pre-pandemic baseline levels the strong association between relative mobility and socio-economic factors has stayed consistent in some cases and changed in others, exacerbating the impacts of Covid-19 along socio-economic gradients. Our findings have important implications for understanding and managing the socio-economic inequities that have arisen, and persist, as a result of the Covid-19 pandemic.

2. Methods

2.1. Network mobility data

We used mobile-phone network mobility data from TELUS communications Inc. through their Insights platform which is a privacy preserving system for analyzing mass-mobility patterns within Canada. We focused our analysis on the province of Ontario, Canada of which approximately 2.65 million devices are included in our analysis. The TELUS data comprise of connections to the cellular tower network over time. There are approximately 90,000 unique receivers within the cellular tower network in Ontario. As a device is moved, it will switch its connection from receiver to receiver (termed handovers), generally taking the most proximal receiver. For each connection, the data contain the start and end time of the connection, along with the geographical coordinates associated with that receiver. From the sequences of receiver connections, we can determine a coarse estimate of movement (i.e., which receivers each device were connected to).

2.2. Estimating home neighbourhood

To estimate the home neighbourhood of each device, we used the network data to identify a proximal subset of network tower receivers associated with the largest cumulative dwell time of that device. This subset of receivers is termed the cluster of home neighbourhood receivers. In the urban areas, these clusters can consist of many receivers, while in the rural areas of only a few. For each device we then computed the weighted average of the geographical coordinates of the cluster of home neighbourhood receivers, where the weights were established proportional to the dwell time at each receiver. We repeated the above process of identifying a cluster of home neighbourhood receivers for each month from January to December 2020 to capture potential (month-by-month) changes.

2.3. Mobility indices

We calculated three indices of device mobility for each device for every day between January and December 2020 to capture changes over time in three unique elements of human mobility (Fillekes et al., 2019): radius of gyration, which can be considered a measure of an individuals geographical range of mobility (Xu, Belyi, Bojic, & Ratti, 2018), time away from home neighbourhood which we associate with working (and partaking in other activities) outside the home (Fillekes et al., 2019), and Shannon’s diversity index applied to the number of places visited which we associate with the complexity (number and variety) of activity places (Perchoux et al., 2014). Each mobility index that we computed has the interpretation that higher values are associated with greater mobility activity associated with that aspect of mobility.

2.4. Geographical range

To measure a device’s geographical range of movement, we used a modified version of the classical radius of gyration (ROG) statistic used in previous mobility studies (Gonzalez, Hidalgo, Barabasi, & Barabasi, 2008; Lee et al., 2021). The classical measure of ROG is calculated as the square root of the sum of the squared deviance (i.e., distance) from the mean center of all observed locations (eq. 1)

\[
ROG_{\text{classical}} = \sqrt{\frac{\sum_i d_{ij}^2}{n}}
\]

Where \(d_{ij}\) is the distance from a given tower location (i) to the mean location (\(\mu\)). We modified the classical ROG measure by replacing the mean center with the point estimate of the home neighbourhood cluster (eq. 2)

\[
ROG = \sqrt{\frac{\sum_i (d_{ij} - \mu_i)^2}{n}}
\]

Where \(d_{ij}\) is the distance from a given location (i) to the home neighbourhood location point estimate. We limit the calculation to only use those locations with a dwell time > 10 min in the calculation of the ROG (in an effort to focus on places that might be associated with a stop or activity). We also remove the cell tower receivers in the home neighbourhood cluster from the calculation. The result of these calculations is that the statistic measures a device’s geographical range of movement relative to its home neighbourhood location.

2.5. Time outside the home neighbourhood

We calculated the time outside the home neighbourhood (TOH) for each device. The calculation simply involved summing the dwell times of all connections associated with receivers that were not part of the home neighbourhood cluster (eq. 3) for each day.

\[
TOH = \sum_{i=1}^{k} dt_i \forall i \notin HC
\]

Where \(dt_i\) is the daily cumulative dwell time at cell tower receiver \(i\), \(k\) is the number of cell tower receivers a device connects to, and HC is the list of cell tower receivers in the home neighbourhood cluster. Because a device’s home neighbourhood cluster usually includes multiple receivers which together their catchment includes a much larger geographical range, this statistic more appropriately represents the time outside of the home neighbourhood area (defined by the catchment of the cell tower receivers in the home cluster), rather than strictly speaking the home. In urban areas the spread of receivers in the home neighbourhood cluster is likely to comprise a much smaller geographical area than in rural contexts.

2.6. Diversity of places visited

To calculate the diversity of places visited we used the Shannon diversity index (DIV; eq. 4)

\[
DIV = -\sum_i p_i \ln p_i
\]
Where \( p_i \) is the proportional dwell time associated with cell tower receiver \( i \) on a given day (\( p_i = \frac{dt_i}{\sum dt_i} \)) and \( k \) is the number of cell tower receivers a device connects to. DIV values near zero indicate low diversity and with higher values associated with increasing levels of diversity. As with the RoG statistic we only included those receivers where the dwell time was >10 min as an approximate measure of only those places associated with stops.

### 2.7. Aggregation procedure

We determined the aggregate dissemination area (ADA) census geographical unit containing each device’s cluster of home neighbourhood receivers geographical estimate. ADAs roughly represent neighbourhoods in urban areas and are aggregated from smaller census units to contain a population between 5000 and 15,000 (2). ADAs are therefore a suitable geographical scale for the analysis of mobility because they are geographically small enough to study geographical variation throughout cities, but large enough to contain a sufficient population for our privacy preserving aggregation procedures. We associated each device’s daily mobility statistics with the ADA associated with its home neighbourhood for each month. When aggregating the mobility statistics to the ADA regions we calculated the mean mobility statistics of all devices within an ADA for every day. We found that the RoG measure showed a high degree of positive skew, while the TOH, and DIV statistics did not appear to be highly skewed. In subsequent analysis, we only analyzed those ADAs that met the criteria of having a minimum of 100 devices for each of the 366 days in our study period (January to December 2020), of which we kept \( n = 1515 \) out of a possible 1557 ADAs.

### 2.8. Comparing to baseline levels

We modelled changes in mobility relative to a pre-pandemic baseline defined as the month of February 2020. We also accounted for day-of-week variability in our baseline. We took the mean value, for each mobility measure, for each day of the week, during February as the baseline. We chose February 2020 as a baseline as it was prior to the onset of most cases in Ontario (only four cases were recorded at the end of February 2020) and no government restrictions on local behaviour occurred until March. We then calculate change in mobility as the ratio of the observed aggregated daily mobility values for each region relative to the baseline value (accounting for day of the week) and we use 100 as the reference mobility value to make interpretations straightforward. We show the daily measures of each of the mobility statistics to demonstrate the day-of-week pattern that can be observed and to demonstrate the fine-scale change in mobility patterns throughout 2020.

To reduce noise associated with daily differences and to smooth out variability in mobility associated with the day of the week we computed weekly averages from the daily mobility indicators beginning on January 1, 2020 and extending to December 29, 2020.

### 2.9. Socio-economic indicators

We included three socio-economic indicators in our statistical analysis to study how changes in mobility were associated with different socio-economic circumstances. We used the Canadian Index of Multiple Deprivation (Government of Canada, S.C, 2019; Matheson, Dunn, Smith, Moineddin, & Glazier, 2012) which is composed of four indicators labelled as economic dependency, ethno-cultural composition, residential instability, and situational vulnerability. We did not include situational vulnerability in subsequent analysis as (for Ontario) it was very strongly correlated with economic dependency. Economic dependency is a composite measure that includes elements related to population age structure, work-force participation, and dependency on social assistance programmes; higher values indicate higher economic dependency or more economically deprived areas. Ethno-cultural composition is a composite measure that includes elements related to identifying as a visible minority, immigration status and place of birth outside Canada, and knowledge of official languages (i.e., French or English). The residential instability index is a composite measure that includes information on housing type, home ownership, residential mobility, and population living arrangements (i.e., marital status and solitary living); with higher values indicative of areas where the local population has less stable residential situations. Along with these three socio-economic indicators, we also chose to include population density, which can be used as a measure of the relative urbanness of different ADAs. Of the 1515 ADAs that had sufficient mobility data, 1509 had complete socio-economic indicator data. Thus, our statistical analysis is focused on the mobility statistics for 52 weeks for each of the 1509 ADAs in our dataset. We tested for correlations between the chosen socio-economic indicators, and found low correlations between the economic dependency index, ethno-cultural index, and residential instability index (Pearson’s \( r \leq 0.5 \)). Population density was moderately correlated with the ethno-cultural index (Pearson’s \( r = 0.58 \)). We decided that it was important to keep population density in the model to help further explain differences between urban and rural areas, and because that although the correlation was moderate, the relationship was non-linear (see Supplementary Material).

### 2.10. Sample characteristics

In studies using big mobility datasets it is important to consider how well the sample reflects the population of interest (here the province of Ontario). In this case, the data contains no information on the individual characteristics associated with devices. Prior to releasing the data TELUS remove from the sample any devices known to be associated with a customer <18 years of age. Therefore, the sample principally relates to devices associated with customers aged 18+ and the patterns therein reflect those associated with mobile phone customers in Canada more broadly where ~97% of those aged 18–34 subscribe to a mobile phone plan, dropping down to only 70% in those aged 65+ (CRTC, 2019), thus there is likely a small bias in the data towards younger populations. In terms of socio-economic characteristics of the sample, we again have no information at the device level. To explore how the sample was distributed across the economic deprivation index, we categorized ADAs into quintiles (5 equally sized groups) and compared the sample proportion with census population proportion across the five groups (Table 1). We repeated this same procedure for the three other socio-economic variables we include in our analysis. We found that the sample was relatively unbiased (at the aggregate level) but we found small differences between the sample and the underlying population in terms of the ethno-cultural index (the sample proportion was slightly larger in areas with lower ethno-cultural index values) and similarly with population density (the sample proportion was slightly larger in areas with lower population densities).

### 2.11. Statistical analysis

#### 2.11.1. Exploratory data analysis

We first explore the temporal change in each of the three mobility statistics (ROG, TOH, and DIV) over time as it relates to key dates in the pandemic response in Ontario Canada. Here we explore the daily mobility statistics, which are useful for highlighting the day-of-week patterns that exist within the data. We examine the spatial distribution of the three mobility statistics averaged over a 7-day period associated with the first wave (7-day period beginning April 1, 2020) and the second wave (7-day period beginning November 11, 2020). To explore the relationship between each mobility statistic and each socio-economic variable we computed Pearson’s correlation coefficient for each day in 2020 and plotted how the correlation between each mobility statistic and the socio-economic variables changed over time.
2.11.2. Spatial SUR model

The pairwise correlation analysis between the daily mobility statistics and the socio-economic indicators is useful for exploring the data, but are limited from a statistical standpoint in two ways. First, it does not consider the spatial autocorrelation that is present in the data, and thus will lead to an underestimation of the variance, which will mean that the confidence intervals associated with the correlation coefficient will be too narrow (Haining, 1991). Further, our analysis looks at each independent variable separately and a better approach would be a multivariate setting which will allow us to simultaneously control for the effect of the other covariates.

Seemingly unrelated regression (SUR) models are multivariate regression models that are commonly used when the data exhibits a cross-sectional structure across T time periods. The origins of the SUR approach are described by Zellner (1962) and the approach is used widely in econometrics. The classical SUR model which does not include any spatial effects is effectively a system (or stack) of equations which can be represented as:

\[
\begin{bmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_T
\end{bmatrix} = \begin{bmatrix}
X_1 & 0 & \ldots & 0 \\
0 & X_2 & \ldots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \ldots & X_T & 0
\end{bmatrix} \begin{bmatrix}
\beta_1 \\
\beta_2 \\
\vdots \\
\beta_T
\end{bmatrix} + \begin{bmatrix}
\epsilon_1 \\
\epsilon_2 \\
\vdots \\
\epsilon_T
\end{bmatrix} \tag{1}
\]

Where \(Y_i = (y_{1i}, \ldots, y_{Ni})\) is a \(N \times 1\) vector representing the dependent variable (in our case the four socio-economic indicators), with associated \(K \times T\) vector of coefficients \(\beta_i (\beta_i = \beta_{i1}, \ldots, \beta_{iN})\); and \(\epsilon_i\) is the vector of \(n\) residuals (\(\epsilon_i = \epsilon_{i1}, \ldots, \epsilon_{iN}\)) associated with the spatial ADA regions. A key feature of SUR models is that there is temporal dependence structure among the residual vectors described as:

\[
E[\epsilon_t \epsilon_t'] = \Sigma, \tag{2}
\]

Where \(\Sigma_{tt'}\) is the error covariance between the equation at time \(t\) and the equation at time \(t'\), typically collected into a variance-covariance matrix \(\Sigma = [\sigma_{tt'}]\). The above formulation is flexible in that it allows for the coefficients to vary over time to capture the temporal dynamics associated with a changing effect of the independent variables over time. Other formulations are possible, for example it can be modified to fix coefficients associated with certain independent variables over time.

Anselin (1988) describes the extension of SUR models to spatial data that are applicable when the residuals of the classical SUR model are found to violate the assumption of spatial independence (López, Mur, & Angulo, 2014). In application, spatial SUR models typically focus on cross-sectional data in regional units (such as provinces or counties) where the response variable varies across T time units and the predictors may (or may not) vary across time, emphasizing the widespread applicability of these models for spatial-temporal analysis of geographical data (Gallo & Dall’erba, 2006; Rey & Montouri, 1999).

Table 1

| Quintile | Economic Deprivation Samp. | Economic Deprivation Pop. | Ethno-Cultural Index Samp. | Ethno-Cultural Index Pop. | Residential Instability Samp. | Residential Instability Pop. | Population Density Samp. | Population Density Pop. |
|----------|-----------------------------|---------------------------|---------------------------|--------------------------|------------------------------|----------------------------|--------------------------|-------------------------|
| Low 1    | 0.22                        | 0.22                      | 0.19                      | 0.15                     | 0.22                         | 0.19                       | 0.21                     | 0.16                    |
| 2        | 0.20                        | 0.21                      | 0.24                      | 0.20                     | 0.22                         | 0.21                       | 0.26                     | 0.20                    |
| 3        | 0.20                        | 0.20                      | 0.22                      | 0.21                     | 0.20                         | 0.19                       | 0.21                     | 0.21                    |
| 4        | 0.18                        | 0.18                      | 0.20                      | 0.22                     | 0.16                         | 0.18                       | 0.17                     | 0.21                    |
| High 5   | 0.20                        | 0.19                      | 0.16                      | 0.22                     | 0.20                         | 0.23                       | 0.15                     | 0.23                    |

In this paper, we test three different spatial-SUR models where the response variable are the aggregate relative mobility statistics (radius of gyration, ROG; time outside of the home neighbourhood, TOH; and diversity of places visited, DIV). The model uses weekly averages of these mobility statistics across the T = 52 weeks of the year 2020, and the geographical units are the N = 1509 ADA regions in Ontario, Canada. We include \(K = 4\) independent predictors, which are the socio-economic indicators associated with economic dependency, ethno-cultural index, residential instability and population density. It is important to note here that none of our predictors (independent variables) vary over time as they are taken from the most recent census (in 2016).

We first look at the overall and time-varying coefficient-of-determination (R²) values to comment on the overall and relative levels of goodness-of-fit of the three models. Then we investigate how the spatial autocorrelation parameter (rho; \(\rho_t\)) varies over time. Finally, we investigate how the \(\beta\) coefficients vary over time, and between the different models, to better understand how the temporal changes in relative mobility captured by our data over the course of the pandemic related to socio-economic factors. For the model coefficients we plot point estimates (the \(\beta\) estimate along with the 95% confidence interval) against time to study how these relationships changed over time.

3. Results

3.1. Exploratory data analysis

Changes to daily mobility in Ontario, Canada varied greatly between three different measures (Fig. 1). Overall, the ROG measure showed the greatest percent decrease between baseline levels and the peak of the lockdown in April, with the provincial average reaching levels of approximately 50% of baseline condition (Fig. 1b). ROG also showed the highest peaks in the summer, relative to baseline levels reaching on average 100% increase over baseline levels in early August. The ROG indicator is a measure of the range of mobility, and therefore increases in mobility associated with summer tourism activities may be driving the observed pattern during the summer months. The measure TOH is the

\[
\mu_t = \rho_t W + \epsilon_t, \tag{4}
\]

Where \(\mu_t\) is a vector of error terms for each time point \(t\), which is the sum of a spatially structured term and a temporally structured term (\(\epsilon_t\)). The spatial structure in the error terms is modelled via the spatial autocorrelation parameter \(\rho_t\) and is applied to the spatial weights matrix \(W\) using the classical queens-case definition of spatial neighbours. The SUR-SEM model was fit using maximum likelihood (López et al., 2014) as implemented in the spsur package in R (Angulo, Lopez, Minguez, & Mur, 2020).
only mobility indicator to remain (on average) below baseline levels into the second wave (Fig. 1c) which in Ontario, Canada began in mid-August (Fig. 1a), is suggestive of widespread shifts to home-based working throughout the province. However, TOH was not as extreme as ROG with the provincial average reaching its lowest point (approximately 50% of baseline levels) in mid-April; and rebounding to approximately 80–90% of baseline levels throughout the second wave suggesting that this measure is capturing a different property of mobility.

Fig. 1. a) Daily new cases of Covid-19 in Ontario, Canada. Daily mobility levels in Ontario, Canada for aggregate dissemination areas (ADAs; \(n = 1509\)) relative to baseline (February 2020) levels for three different measures of mobility: b) radius of gyration (RoG), c) time outside the home (TOH), and d) diversity of places visited (DIV). Each ADA is represented as a grey line highlighting geographical variability in the mobility response relative to the daily provincial average (in black). The coral ribbon shows the middle 95% range of data (contained between the 2.5 and 97.5 percentiles). The vertical lines show key dates in the lifting of provincial restrictions associated with the pandemic and subsequently the re-implementation of lockdowns. In the second wave, two regions (Toronto and Peel) were placed under lockdown ahead of the rest of the province.
behaviour. The TOH indicator showed a strong weekday-weekend pattern, which again is likely an indicator of changes in home-based working throughout the pandemic. The DIV indicator is a measure of the diversity of places or activities associated with a device and on average it was the most moderate, i.e., after the lockdown the median DIV value decreased only by about 20% to 80% of baseline levels (Fig. 1d). The DIV indicator rebounded to approximately baseline levels after the province had moved into Stage 3 (most open stage) of the lockdown (Fig. 1d), and only dipped again after the onset of new restrictions in late November and December of 2020. We see that however that the range of DIV values varied greatly later in 2020 especially when compared to the pre-lockdown periods (e.g., Jan 2020) which is suggestive of substantial geographical variation in the DIV indicator. 

There was significant geographical variation in the mobility response relative to baseline levels into the second wave as demonstrated by the increasing variation around the mean observed in Fig. 1. Mobility indicators during the peak of the lockdown (7-day period of April 1 to 8) showed a high degree of change in relative mobility during the first wave for each of the three indicators (see Fig. 2 for the ROG index and the supplementary material for the TOH and DIV maps). We can also see that these patterns differ slightly into the second wave (7-day period November 11 to 18; see Fig. 3 for the ROG index and the supplementary material for the TOH and DIV maps). A complete viewing of the geographical variation in relative mobility week-by-week can be seen on the Ontario COVID-19 mobility dashboard: https://geospatial.uwo.ca/mobility.html.

All three measures of mobility were positively correlated with the economic deprivation index in the immediate time period following the lockdown and continuing into the second wave with the ROG index showing the most variability in this relationship (Fig. 4a). The economic deprivation index is inversely proportional with economic standing, and thus higher values are associated with more economically deprived areas. The positive correlation here therefore is further evidence that higher mobility levels were associated with more socially disadvantaged areas. The ethno-cultural index was negatively correlated with TOH immediately following the lockdown, but not associated with DIV until later into the pandemic, and was negatively associated with changes in mobility thereafter (Fig. 4b). The ethno-cultural index is scaled in a way that areas with higher index scores are areas with more ethnic diversity, therefore the negative associations between the ethno-cultural index and TOH and DIV are reflective of areas with more ethno-cultural diversity showing greater reductions in mobility. In the first wave of the pandemic, the residential instability index was positively associated with DIV, but later in the pandemic it was negatively associated with DIV (Fig. 4c). Residential instability showed low correlation with ROG and TOH, throughout the pandemic. The changing sign of residential instability associated with DIV is interesting as it suggests different patterns of mobility over time being captured by this variable. Finally, we see that population density showed a similar association pattern with the mobility measures as the ethno-cultural index, which highlights that in this part of Canada, more ethnically diverse areas are typically in more urban environments (Fig. 4c). We further explore these relationships in a multivariate setting, controlling for present spatial autocorrelation using the spatial SUR-SEM model approach.

3.2. Spatial seemingly unrelated regression (SUR-SEM) models

After fitting the three models, we compare goodness of fit across the models for each of the mobility statistics. The TOH model had the highest overall coefficient of determination – pooled-\( R^2 = 0.80 \), the DIV model had a pooled-\( R^2 = 0.73 \), and the ROG model had the lowest pooled- \( R^2 = 0.60 \). However, it can be seen that \( R^2 \) varied across the 52 weekly periods in the year, with the highest \( R^2 \) values associated with the summer period after the relaxation of lockdown conditions associated with the first wave (Fig. 5a). We can also see that in general the models typically had a higher \( R^2 \) after the onset of the restrictions to curb the spread of Covid-19 in Ontario, Canada occurring in the middle of March (the exception being the first week of January). It is also interesting to note that the general trend in the TOH and DIV models was somewhat similar, but that this differed from the ROG model.

Looking at the pattern of spatial autocorrelation coefficient (rho) over time, we note that the temporal dynamics are similar (but not identical) to the temporal pattern of \( R^2 \) values (Fig. 5b). The parameter rho was relatively constant for the TOH and DIV models (mean rho for
TOH = 0.46; mean rho for DIV = 0.47). The rho values for ROG were overall lower (mean rho for ROG = 0.18), and varied across a wider range [0.036, 0.52]. We found that the spatial autocorrelation coefficient (rho) was significant ($p < 0.05$) for every week for the DIV and TOH model, and for every week in the ROG model except for at the lowest point (in May; $p = 0.13$), indicative of the strong degree of overall spatial autocorrelation present. The spike in the coefficient for rho for the ROG model during the relaxation of the lockdown restrictions in July is interesting. The peak in rho (and similar peak in $R^2$) occurring during the summer months is coincident with a period of higher mobility associated with this time period as captured by ROG (Fig. 1b).

Next, we explore the temporal dynamics of the coefficients associated with each of the covariates for each of the models for ROG, TOH, and DIV. First, we note that our relative measures of mobility were derived with a baseline value of 100 (representing average conditions pre-lockdown during February 2020). Thus, when the intercept is significantly different from 100 it is evidence of changes to overall relative mobility not captured by variation in the socio-economic covariates. The intercept coefficient was typically near 100 in the pre-pandemic phase for each of the three models (Fig. 6), but after the initial lockdown occurred the intercept becomes significantly lower than 100 for each model, reflective of the overall change in daily mobility observed in Fig. 1. We can see in the model for ROG and DIV the temporal evolution of the intercept coefficient goes significantly above the baseline level of 100, however this is not observed in the TOH model. This is interesting given that ROG was the only measure that showed average mobility levels much greater than the baseline in the summer months (Fig. 1).

The economic dependency index showed low levels of association with relative mobility prior to the initial lockdown (Fig. 6). However, we observe that the economic dependency index shows a strong and consistent positive association with all three mobility indicators for the duration of the first and second waves. Further, there is evidence that this strong and positive association began before the onset of the first lockdown in all cases. The ROG model shows a short period of no association (1 week) with economic dependency occurring at the beginning of September, and later shows a negative association during the last week of the year (after the implementation of the second lockdown province wide).

The ethno-cultural index showed low association with mobility statistics in the time preceding the first lockdown along with the early portions of the first lockdown. Specifically, with ROG the ethno-cultural index was showed a small positive association with ROG in the weeks immediately following the first lockdown but then showed no association later in the first lockdown. After the removal of the first lockdown restrictions, we see a strong negative association between the ethno-cultural index and the ROG measure for the remainder of the summer months. Into the second wave we see a strong positive association between the ethno-cultural index and ROG. The pattern is different for the other mobility metrics, as the ethno-cultural index showed low association with the TOH and DIV measures in the early parts of the first lockdown, and then was subsequently negatively associated with these relative mobility measures into the second wave.

With the residential instability index, we see a completely different pattern of association with the relative mobility measures. First, we see that coefficient for residential instability shows very low levels of association with ROG throughout the pandemic, the exception being some key periods associated with the beginning of January and late December. We also see that prior to the pandemic RI was negatively associated with ROG, but immediately following the first lockdown, there is a two-week period where the coefficient is positive and significant. RI was positively associated with TOH throughout the first and second wave of the pandemic, but showed no association with DIV, except during the first few weeks after the first lockdown.

The temporal pattern of the coefficient for population density differed between the three different relative mobility measures. Population density was negatively associated with ROG in the first wave, but by the summer showed no association with ROG. In the second wave of the pandemic ROG was negatively associated with population density, with the timing of this period corresponding to the beginning of September. The coefficient for population density in the TOH model showed a very clear and abrupt negative association throughout both the first and second wave, but no association prior to the first lockdown. Finally, the coefficient for population density in the model for DIV
showed a positive association in the first weeks after the first lockdown occurred, but in May exhibited a sharp change and the coefficient was negative throughout the second wave.

A main finding is that the pattern of association between socio-economic variables (economic dependency index, ethno-cultural index, residential instability index, and population density) and mobility varied across the timeline of the pandemic in Ontario, Canada. A second main finding is that the nature of these associations changed depending on mobility is being measured (Table 2). For example, we found that in the 42 weeks that followed the onset of the first lockdown in Ontario, Canada, the ROG, TOH, and DIV mobility measures were all positively associated with economic deprivation (Table 2). Interestingly, the ROG measure had at least one week that there was a positive association and at least one week with a negative association for each of the socio-economic covariates (including the intercept) during this period.

The TOH was found to be significant and positively associated with economic deprivation, and residential instability, and negatively associated with population density for all of the 42 weeks post lockdown. The DIV indicator was similar to that of TOH, with the main difference occurring in the association with the residential instability index, where in the DIV model DIV was not found to be significantly associated with residential instability for 37 of the 42 weeks post-lockdown.

4. Discussion

We used three different measures of mobility in an attempt to capture three fundamentally different aspects of human mobility. Specifically, we implemented measures related to the geographical range of movement, time outside of the home neighbourhood (a measure of daily working patterns), and diversity of places visited (a measure of the
number and variety in activities). The temporal patterns of aggregated relative mobility we identified in Ontario, Canada was associated with rapid decreases in mobility behaviour on the order of magnitude of 50% of pre-pandemic conditions. Other regions across the world experienced similar changes in daily mobility associated with initial lockdown conditions (Gao, Rao, Kang, Liang, & Kruse, 2020; Pepe et al., 2020; Pullano, Valdano, Scarpa, Rubrichi, & Colizza, 2020; Warren & Skillman, 2020). The ROG and TOH measures both reached relative mobility levels 50% of the baseline in mid-April at the peak of the first lockdown, however the provincial average of the DIV measure only reached approximately 80% of baseline levels in the first lockdown. The rapid change in mobility during the first lockdown has been widely reported, and was shown to be correlated with differential rates of Covid-19 in the first wave (Badr et al., 2020; Chang et al., 2020; Xiong et al., 2020), but not in subsequent waves (Gatto et al., 2020). In Ontario, mobility levels (on average) returned to baseline levels in the summer months and stayed at this level into the second wave.

It is important to note that the relative mobility measures are here defined relative to February 2020 levels. This was necessary due to the inability to access data from prior to 2020. However, mobility patterns in Ontario are seasonal, with typical mobility levels much higher in the summer months than during the winter (Abdelgawad, Abdulazim, Abdulhai, Hadayeghi, & Harrett, 2015). The fact that overall mobility was only slightly above or equal to the baseline conditions during these months is indicative that mobility levels in the summer of 2020 were still substantially lower than at the same time in previous years.

We use three different mobility measures derived from cellular devices aggregated (by taking the mean value) across relatively fine geographical units. This approach was warranted for our study of socio-economic associations, because we are interested in studying the different components of mobility. Alternatively, many studies have simply studied mobility defined as origin-destination flows between regions (Gao et al., 2020). This approach typically requires the use of larger aggregation units to capture suitable numbers of flows for both privacy and statistical analysis reasons. Mobility data as flows have also been integral to many different types of models used to estimate future Covid-19 rates (Karatayev, Anand, & Bauch, 2020; Zachreson et al., 2021). The finer-scaled mobility indicators used here (comprising three different indicators of mobility) might provide alternative opportunities for improved predictions of Covid-19 rates in subsequent waves of the pandemic.

Aggregate measures of mobility are necessary when working with large mobility datasets because aggregation serves as a privacy preserving operation when combined with suitable minimum inclusion criteria (e.g., here we only kept ADA regions with a minimum of 100 devices for subsequent analysis). Individual level studies are often combined with detailed individual level survey information which provide rich datasets for analysis of socio-economic factors associated with mobility at the individual level (Guo, Chai, & Kwan, 2020; Helbich et al., 2016; Kwan, 1999; Long & Reuschke, 2021; Schwanen, Kwan, & Ren, 2008). Here we do not have access to detailed individual level data, and thus focus on aggregate patterns, with a large sample of the population. We use a derived estimate of the home neighbourhood location to associate a device-specific mobility pattern with a specific ADA region, and relate aggregate socio-economic conditions to aggregate mobility patterns.

Here we used census-level socio-economic data in combination with aggregate mobility statistics to study the relationships between mobility and socio-economic conditions. Therefore, because of the aggregation process, and the use of administrative geographical units, our results are subject to the modifiable areal unit problem (MAUP) (Openshaw, 1984). The choice of the ADA level scale of analysis was based on the fact that this was the smallest census unit that still allowed us to include most regions in the final analysis (1509 of 1557 ADAs were included) based on our n = 100 devices inclusion criteria. Smaller census units (i.e., dissemination areas; DAs) are much smaller than ADAs both in population (DAs = 400–700 population vs ADAs = 5000–15,000 population), but also in geographical scale. Dissemination areas were deemed problematic both from the perspective of reaching the inclusion criteria of 100 devices, but also because the network mobility data we used is based on the spatial distribution of cell tower receivers, which are relatively sparse in areas with lower population density. Importantly, the use aggregate dissemination areas (ADAs) represent a finer geographical scale of analysis than what has presently been used to study the relationship between socio-economic variables and relative mobility patterns during Covid-19. Thus, our analysis provides novel insights about the relationships between aggregated relative mobility during Covid-19 and socio-economic factors at geographical scales that have previously not been explored.

The associations between the aggregate level socio-economic variables and relative mobility statistics were explored using both pairwise correlations and the multivariate spatial seemingly unrelated regression models (SUR-SEM). Here we focus discussion on the results from the SUR-SEM models. It is important to highlight here that both the correlation analysis and the SUR-SEM models only provide evidence of associations between relative mobility measures and socio-economic indicators (and how these changed over time) and we cannot infer causation based on these analyses alone.

In general, the three SUR-SEM models (one for each relative mobility measure) showed relatively good overall fit as defined by $R^2$, but we also found that in general the fit of the models varied over time. The $R^2$ of all three models increased around the time of the first lockdown and peaked in the summer months when mobility levels were highest. One potential explanation for this is that Covid-19 immediately caused changes in mobility stratified across socio-economic gradients immediately at the onset of the first lockdown (Dasgupta et al., 2020) which were being captured by our model. These relationships then became further
magnified during summer months as people settled into new lockdown routines and could selectively participate in discretionary travel. It is interesting then that the ROG measure had the lowest overall $R^2$, and the differences in $R^2$ between the ROG model and the other two grew larger following the onset of the pandemic because ROG is one of the most widely used measures of mobility in studies using large mobile phone datasets.

The strong positive relationship between economic dependency and the three relative measures of mobility further highlights how the pandemic has led to what Huang et al. (2021) refer to as the luxury of social distancing. This is an interesting finding because mobility levels, in particular the measures of the geographical range of mobility – like ROG – would typically be negatively associated with economic dependency (Morency, Paez, Roorda, Mercado, & Farber, 2011). Likewise, previous research has shown that those with higher socio-economic standing to also have more complex activity patterns (Roorda, Paez, Morency, Mercado, & Farber, 2010) which was not reflected in the DIV model. The clear change in the coefficient for economic dependency in the TOH model at the onset of the lockdown is striking, and reflects previous findings in the United States (Jay et al., 2020) and supports the notion that those who are more highly educated and in more skill-based sectors were most easily able to transition to home-based working.

Fig. 6. Temporally varying ($\beta$) coefficients for the spatial seemingly unrelated regression (SUR-SEM) models for radius of gyration (ROG; left), time outside the home neighbourhood (TOH; middle), and diversity of places visited (DIV; right). The temporally varying coefficients are for the intercept term, and the socio-economic indicators for economic dependency (ED), the ethno-cultural index (EC), residential instability (RI), and population density (PD). The plots show the point estimate as a filled circle and the 95% CI of the point estimate as a vertical line; black is used to indicate those CI’s that do not include the expected value (100 for the intercept and 0 for the others), grey is used to show the CI’s that do include the expected value. The red vertical line indicates the onset of the first lockdown, the green vertical line is when the province of Ontario moved to the least restrictive stage during the summer and the orange vertical line refers to the start of the second lockdown (beginning in the Toronto area). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Table 2
For each of the 42 weeks after the initiation of the first lockdown in Ontario, Canada (March 16, 2020) we count the number of times the β coefficient showed negative (−), no significant (No), or positive (+) association between the three mobility indicators used as dependent variables (ROG – radius of gyration, TOH – time outside home neighbourhood, and DIV – diversity of places visited) and the socio-economic covariates included in the SUR-SEM model (ED – economic deprivation index, EC – ethno-cultural index, RI – residential instability index, PD – population density).

|        | ROG  | TOH  | DIV  |
|--------|------|------|------|
| Intercept | 24   | 5    | 13   |
| ED      | 12   | 10   | 3    |
| EC      | 4    | 3    | 2    |
| RI      | 1    | 1    | 1    |
| PD      | 30   | 9    | 3    |

(Bartik, Cullen, Glaeser, Luca, & Stanton, 2020) and it appears this pattern as continued throughout the second wave of Covid-19 in Ontario, Canada. The economic deprivation index was drawn from the 2016 census data, however in response to Covid-19, the Canadian government implemented an emergency social assistance programme at a scale previously not seen before. Individuals who lost their job due to pandemic-related closures were able to apply for a $2000 CAD (~$1600 USD) per month benefit. The strong association between economic dependency and mobility, in particular TOH, is thus even more striking given that people in lower income areas were more likely to have accessed the Government Covid-19 benefit programme (Achou et al., 2020). Thus, the pattern of association between mobility and economic deprivation is further evidence to support the disproportionate impact Covid-19 has had on mobility levels across economic gradients, likely shaped in large part by those essential workers at the lower end of the income scale.

The impacts of Covid-19 infections are widely reported to fall disproportionately on different ethnic groups and those with lower educational attainment levels (Sundaram et al., 2021). However, less has been reported on how changes in relative mobility in response to pandemic related lockdowns are related to the ethnic diversity of regions. Here we found the relationship between the ethno-cultural index and relative mobility varied across time and between the three different mobility measures. The pattern of the strong negative and significant coefficient for EC in the ROG model during the summer months provides evidence that areas with higher ethno-cultural index values did not engage in travel that resulted in large displacements from their home neighbourhood during these summer months. People that live in areas with a higher ethno-cultural index values are more likely to be recent immigrants and/or not speak one of English or French. In Ontario, many regions with higher ethno-cultural index values are found in the urban and sub-urban regions around Toronto. The strong negative association with TOH here is evidence that areas with a higher ethno-cultural index were found to spend more time in the home-neighbourhood than other regions, which contrasts with patterns of reported Covid-19 in Ontario which were positively associated with ethno-cultural indicators (van Ingen et al., 2021). Thus, these results further highlight how different mobility indicators may lead to different associations with socio-economic factors.

The residential instability index relates to building characteristics, household composition, and residential mobility. The coefficient for the residential instability index was positive in the ROG model at key junctures in the calendar year, specifically the week following New Years, the week immediately following the implementation of the first lockdown, during the summer months after lockdown restrictions eased, and at Christmas coinciding with the second lockdown. A likely explanation for this is that this variable is capturing areas with a younger population and/or students who are engaged in long distance residential changes during these key periods, which would be captured by the ROG measure. Similarly, we see that the coefficient for residential instability index in the TOH model was consistently positive following the first lockdown, which reflects what we know about the employment situations of many younger populations being disproportionately expected to work outside the home during the pandemic as essential workers. However, we see no association between residential instability and DIV throughout the pandemic, suggesting that the characteristics of residential situation where not associated with this property of mobility.

We found that areas with higher population density were associated with lower ROG in both the first and second wave, but coincidentally higher ROG during the summer as restrictions were eased. We also found that areas with higher population densities were associated with lower TOH and DIV measures, but interestingly immediately following the lockdown there was a brief period where the coefficient for population density indicates a positive association with DIV. The positive coefficients observed during the summer months in the ROG model are consistent with reports of residents from denser areas, particularly in Toronto, leaving the city to visit cottages and/or participate in local tourism to escape the most densely populated part of the province throughout the first wave of the pandemic (Andrew-Gee & Bula, 2020; Goldfinger, 2020).

5. Conclusions
Our analysis demonstrates first and foremost the ongoing inequities associated with Covid-19 in terms of differential impacts on relative mobility patterns; these inequities fall clearly across socio-economic gradients. Specifically, here we explore how changes in mobility were associated with economic dependency, ethnic/cultural diversity, residential instability, and population density. Further, we demonstrate that the associations between socio-economic variables and relative mobility change over the course of the pandemic, in some cases completely switching in directionality. This finding has important implications for understanding how the pandemic, and resulting imposed restrictions, have differentially impacted different populations at different times. Finally, the results from our analysis emphasize the fact that how mobility (or relative mobility) is measured will influence the relationships between socio-economic variables and mobility; an outcome that to this point has been largely ignored in the race to study changes in human mobility patterns in response to Covid-19.

Acknowledgements
This work was supported by the TELUS Data for Good Program. We are grateful for assistance from W. Li and J. Bettridge at TELUS who contributed to this project. Funding was provided from a Western University Catalyst Grant to JL. The authors would also like to thank the Associate Editor I. Benenson and the reviewers for their helpful feedback.

Appendix A. Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.compenvurbs.2021.101710.

References
Abdelgawad, H., Abdulazim, T., Abdullahi, B., Hadaveghi, A., & Harrett, W. (2015). Data imputation and nested seasonality time series modelling for permanent data collection stations: Methodology and application to Ontario. Can. J. Civ. Eng., https://doi.org/10.1139/cjce-2014-0087
Achou, B., Boiteclair, D., D’Astous, P., Fonseca, R., Glenzer, F., & Michaud, P.-C. (2020). Snapshot of households that received the Canada emergency response benefit and paths for further investigation. Cirano, Survey of Household Finances in a Time of Pandemic. & Andrew-Gee, E., & Bula, F. (2020). Urbanites run for country homes, cottages amid coronavirus outbreak, creating tensions with year-round residents. Globe Mail.

J.A. Long and C. Ren
Computers, Environment and Urban Systems 91 (2022) 101710
Zachreson, C., Mitchell, L., Lydeamore, M. J., Rebuli, N., Tomko, M., & Geard, N. (2021). Risk mapping for COVID-19 outbreaks in Australia using mobility data. *J. R. Soc. Interface, 18*, 20200657. https://doi.org/10.1098/rsif.2020.0657

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation Bias. *J. Am. Stat. Assoc.*, 57, 348–368. https://doi.org/10.1080/01621459.1962.10480664

Zhang, J., Litvinova, M., Liang, Y., Wang, Y., Wang, W., Zhao, S., Wu, Q., Merler, S., Viboud, C., Vespignani, A., Ajelli, M., & Yu, H. (2020). Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science*, 368, 1481–1486. https://doi.org/10.1126/science.abb8001