ECARES European Center for Advanced Research in Economics and Statistics ULB

The Spread of COVID-19 in Belgium: a municipality-level analysis

Philip Verwimp
ECARES, Université libre de Bruxelles
Email: philip.verwimp@ulb.ac.be

COVID-19 Demography Conference, Vienna, Nov 30-Dec 1, 2020

introduction

- I want to investigate the spread of the virus in Belgim by looking at three variables of intrest:
  - the onset of the epidemic, meaning the first infections
  - the intensity of the epidemic, meaning the infection rate
  - the growth of the epidemic over time

- And I do that at the level of the municipality (N=581) in order to analyse the socio-economic and demographic correlates of onset, intensity and growth

Data

- Provided by Sciensano, the public health agency at the federal level
- Daily updates of number of infections for all municupalities
- Period under investigation is March 1 till May 4, testing-policy was uniform across the country in this period, coincides with first wave of the epidemic and first lockdown in Belgium, with nationwide uniform rules of enforcement
- Completed with socio-economic and demographic variables that are readily available for all Belgian municipalities from government data sources

Relevance of the municipality level

- On average a Belgian municipality is small, 5 to 10 km to drive through
- Large part of life occurs within it: primary and seconday schooling, sport activities, socio-cultural associations, smalles administrative and political entity
- A team around Piet Maes (Rega Institute, KU Leuven) has sequenced 250 SARS-CoV-2 genomes in Belgium. His spatially-explicit phylogeographic analyses allow to track the dispersal velocity of viral lineages. The dispersal velocity of viral lineages was 5.4 km a day before the lockdown and 1.2 km a day in the first few weeks of the lockdown. (see Dellicour et al, 2020).

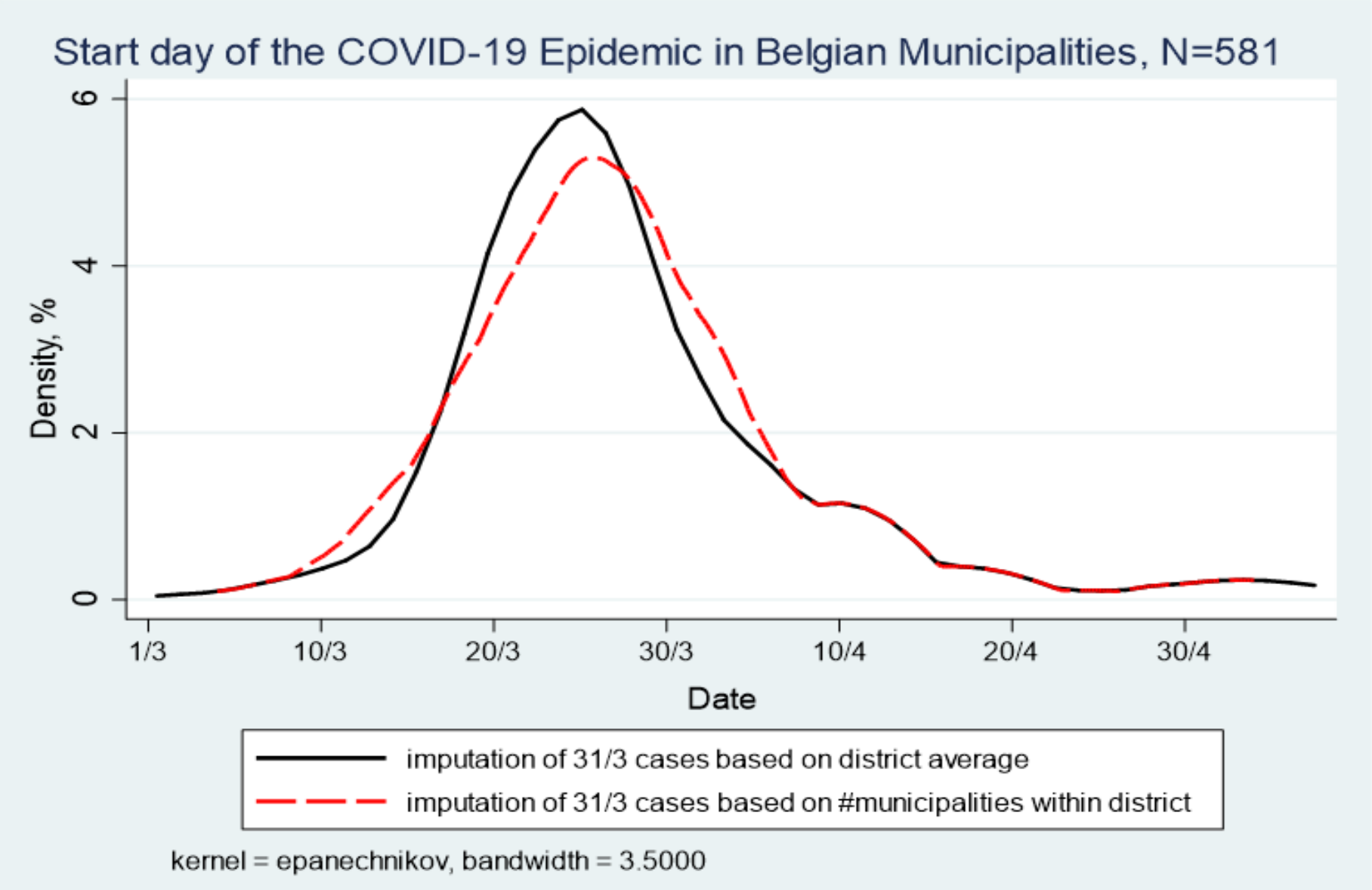
Preview of the findings

- In March, COVID-19 spread faster in larger, more densely populated, higher income municipalities with more elderly people and a larger share of the elderly population residing in care homes.
- Municipalities which were more exposed to migration, foreign travel for business, leisure or family affairs were affected earlier on in the epidemic.
- Richer municipalities managed to slow down the epidemic in April more compared to poorer ones.
- Income correlates with the infection rate in particular in the Flemish Region whereas the share of foreign nationalities correlates with the infection rate in particular in the Walloon Region.

(1) Onset of the epidemic

- I investigate the start date of the epidemic in each municipality
- Sciensano only published the exact figure of infections on a given day if this figure is larger than 5. It also publishes the cumulative number of infections each day.
- In order to analyse the socio-economic and demographic correlates of the onset I first have to estimate the first day where the number of infections was smaller than 5.
- From the 581 municipalities we have the exact date of the start of the epidemic, defined here as at least 5 confirmed contaminations for 309 municipalities of which 205 reached that threshold in March and 104 in April. For the remaining 272 municipalities Sciensano published the accumulated number of contaminated cases on March 31.
- For these 272 municipalities I inpute the day of onset as the average day at the district level, one administrative unit above the municipality level.

Kernel density estimation of start date



onset

- We derive from this that
  - 4% of all municipalities had their onset before March 15
  - 70% between March 15 and March 30
  - and 26% after March 30

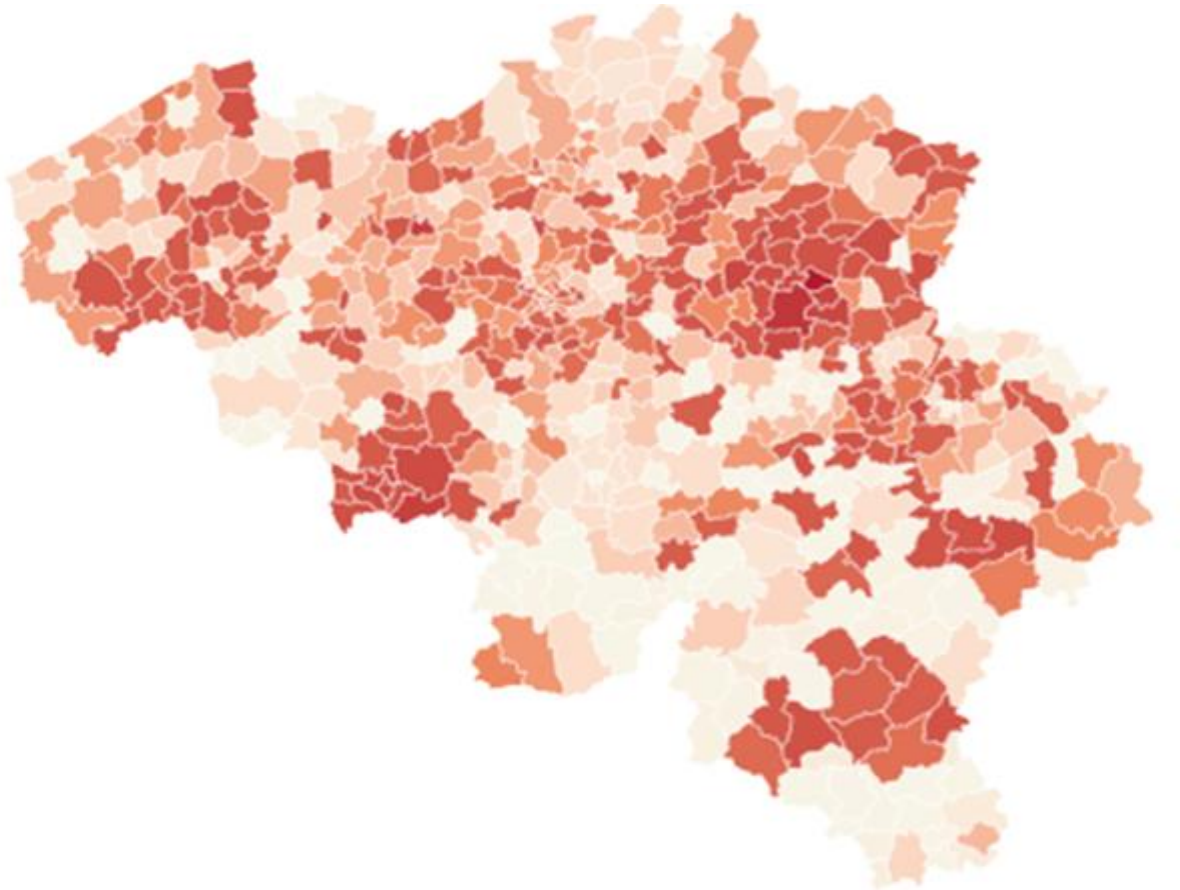
- I can now estimate the correlates of the start date. In the paper I formulate a number of hypotheses that I want to test with the data. These are based on the very first cases that are reported in Belgium, namely from business travel to Wuhan in China and from leisure and ski travel to northern Italy.

Dep. Var.	Start date of the COVID-19 Epidemic, in #days after March 1					
	R1	R2	R3	R4	R5	R6
Indepen. Var.						
Pop.size ('1000)	-.10*** (.03)	-.12*** (.04)	-.12*** (.036)	-.11*** (.036)	-.11*** (.032)	-.10*** (.03)
Population density ('1000)	-.45*** (.17)	-1.17*** (.23)	-1.14*** (.22)	-1.02*** (.22)	-.54** (.23)	-.001*** (.0002)
% of population +65 of age		-.64*** (.13)	-.45*** (.15)	-.37** (.16)	-.48*** (.16)	-.48*** (.16)
% singleton households		.46*** (.10)	.31*** (.12)	.27** (.12)	.30*** (.11)	.39*** (.13)
Income per cap. ('1000)			-.55*** (0.18)	-.56*** (.17)	-.40** (.17)	-.52*** (.19)
% of pop +65 in care homes				-.25* (.13)	-.26* (.13)	-.32** (.13)
Municipality at the border					4.44*** (1.36)	1.44 (1.45)
%Foreign Nationalities					-.20** (0.08)	
% French						.27**(.12)
% German						.10 (.17)
% Dutch						.005(.11)
% Italian						-1.28*** (.19)
% Chinese						-14.3*** 4.26)
Constant	29.91*** (.64)	28.79*** (3.21)	40.42*** (4.80)	41.41*** (4.79)	39.89*** (4.59)	40.36*** (5.21)
N	581	581	581	581	581	581
R2	0.14	0.21	0.23	0.23	0.26	0.30

(2) Intensity

- The number of infections per 1000 inhabitants
- On March 31, a few days before the epidemic reaches it “high point” in the first wave in Belgium, at which the number of new contaminations is lower then the day before
- 12,300 infections at that point
- I regress the infection rate on the same set of socio-demographic and economic variables as in the case of onset

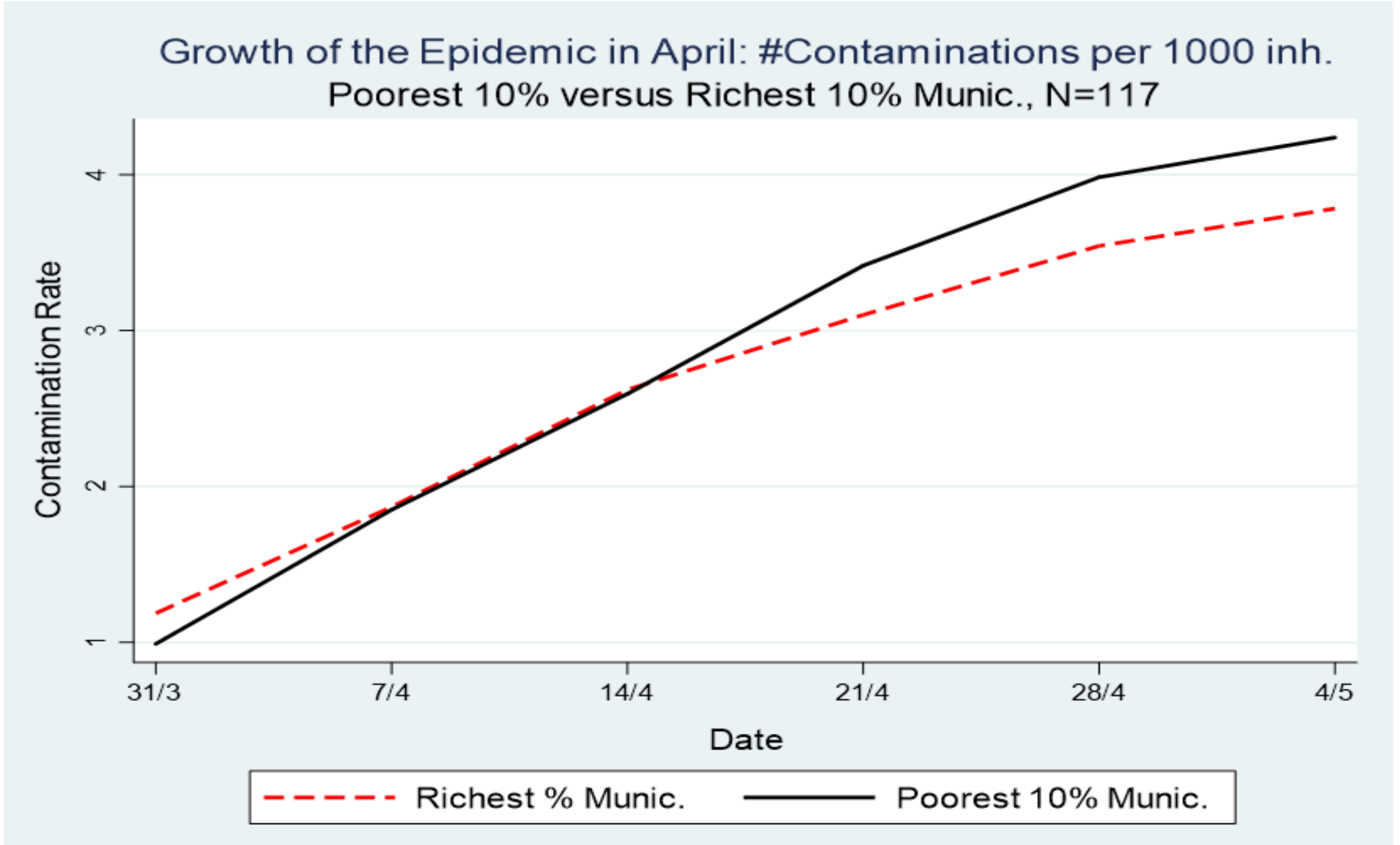
Number of registered infections per 1000 inhabitants on March 31



Dep. Var.	Contamination rate per 1000 inhabitants on March 31						
	R1	R2	R3	R4	R5	R6	R7-IV
Ind. Var.							
Start day of epidemic						-.05***	-.05***
						(.004)	(.012)
Pop density	.06***	.06****	.05***	.03**	.04***	.01	.007
('1000)	(.01)	(.01)	(.01)	(.017)	(.01)	(.01)	(.02)
% pop +65	.041***	.042***	.034***	.039***	.041***	.028**	.026***
	(.011)	(.007)	(.012)	(.012)	(.013)	(.011)	(.010)
% singleton households	-.017***	-.015**	-.013*	-.014**	-.023***	-.024***	-.023***
	(.006)	(.007)	(.007)	(.007)	(.009)	(.008)	(.006)
Income per capita		.006	.006	-.002	-.001	-.02	-.02
		(.013)	(.013)	(.014)	(.015)	(.015)	(.014)
%pop+65 in			.017*	.017*	.018*	-.002	-.0003
care homes			(.010)	(.01)	(.009)	(.01)	(.009)
Border				-.22**	-.06	-.0002	
				(.09)	(.10)	(.09)	
%Foreign Nationalities				.006 (.005)			
% French					-.015 (.014)	-.0008 (.01)	
% German					-.014** (.006)	-.01* (.05)	
% Dutch					-.006 (.008)	-.013** (.007)	
% Italian					.05** (.02)	.-004 (.02)	
% Chinese					.88***(.25)	-.09 (.45)	
Constant	.67***	.54	.49	.60	.73*	3.12***	3.03***
	(.25)	(.36)	(.35)	(.36)	(.40)	(.45)	(.74)
N	581	581	581	581	581	581	581
R2	0.04	0.04	0.04	0.05	0.08	0.36	0.35
Underid. test							23.63***
First stage							Start day
% italian							-1.24***
							(.23)
F-test							29.03***

(3) Growth

- I estimate the growth of the epidemic over time, in particular the change in the weekly, cumulative infection rate accross municipalities
- By the end of the first lockdown Belgium registered 50,000 infections
- The correlation between the dependent and lagged dependent is very high, understandably in an epidemic. In Belgian municipalities during COVID-19, it is as high as 0.80 between March 31 and April 7 and still remains 0.60 between March 31 and May 4.
- Inclusion of lagged dependent variable explains large part of the variation in the model. Results do not change with IV model.
- I will look in particular at the effect of poverty/wealth


Growth of the Epidemic in April: #Contaminations per 1000 inh.
Poorest 10% versus Richest 10% Munic., N=117
Contamination Rate
1
2
3
4
31/3
7/4
14/4
21/4
28/4
4/5
Date
Richest % Munic.
Poorest 10% Munic.

Difference-in-Differences, #registered contaminations per 1000 inhabitants in the course of the month of April, by poorest and richest decile, N==117

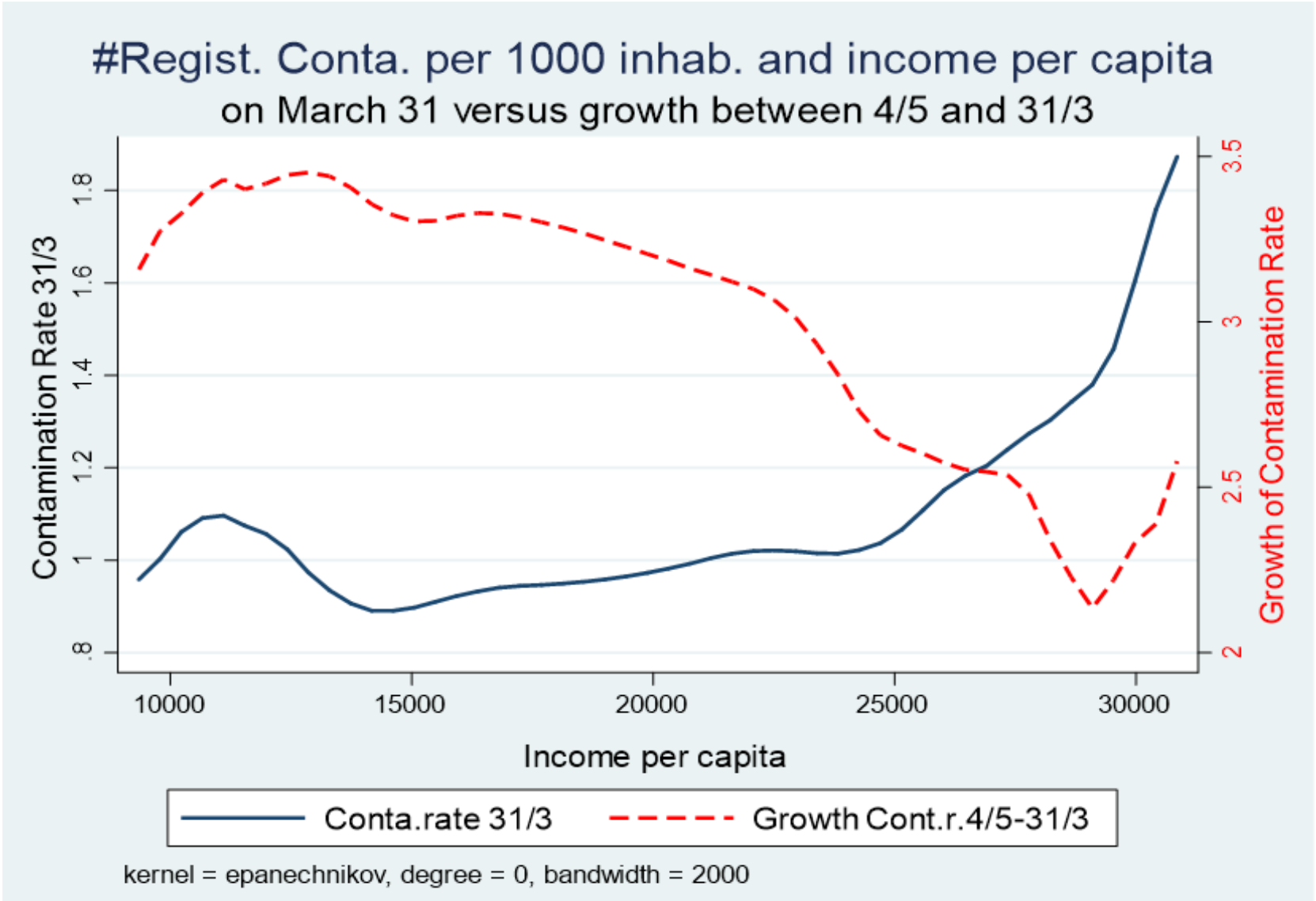
Week	March 31	May 4	Difference
Income decile			
Poorest	0.99	4.24	3.25***
Richest	1.18	3.78	2.60***
Difference	-0.19*	0.46	0.65**

Magnitude of this effect

The difference between 31/3 and 4/5 on the one hand, and between rich and poor municipalities on the other hand, is statistically significant at the 5% level. At the end of the observation period (May 4), this difference amounts to .65 cases per 1000 inhabitants. For a municipality in the poorest decile, of, on average 35,000 inhabitants, this accounts for a difference of 22 more persons infected compared to a municipality in the richest decile. Multiply that figure 58 times for all municipalities in this poorest decile and one obtains a difference of 1,276 more infected persons compared with the richest decile.

Looking at the entire income distribution

- We can also look at the entire income distribution over all municipalities and visualise the effect of income on the growth trajectory of the epidemic by comparing the infection rate on March 31 with that at the end of the observation period (May 4). We notice that the initial disadvantage (early infections) and later advantage (slower growth of the rate of infections) kicks in at around 20,000 euro per capita per year. Around 30% of Belgian municipalities have more than 20,000 euro per capita income in 2018 according to fiscal data.


#Regist. Conta. per 1000 inhab. and income per capita
on March 31 versus growth between 4/5 and 31/3
Contamination Rate 31/3
.8
1
1.2
1.4
1.6
1.8
Growth of Contamination Rate
2
2.5
3
3.5
10000
15000
20000
25000
30000
Income per capita
Conta.rate 31/3
Growth Cont.r.4/5-31/3
kernel = epanechnikov, degree = 0, bandwidth = 2000

Dep.Var	Growth of the Conta. Rate between May 4 and March 31					
	Lag is 6 weeks					
	Log-linear	Linear model				
	R1	R2	R3	R4	R5 - IV	R5 - IV
Indep. Var.	Belgium	Belgium	Flemish Region	Walloon Region	Belgium	Belgium
Lagged Conta.Rate	.035*** (.028)	2.07*** (.16)	2.46*** (.21)	1.63*** (.18)	2.38*** (.68)	2.14*** (.62)
Pop density ('1000)	-.004 (.01)	-.01 (.04)	.23 (.29)	.28 (.35)	-.003 (.037)	-.0003 (.036)
% pop +65	.015* (. 009)	.06* (. 03)	.08 (.05)	.09 (.06)	.07** (.034)	07** (.03)
% singleton households	-.005 (.006)	.017 (.025)	-.05 (.041)	.06 (.04)	-.003 (.019)	-.003 (.019)
Income per capita	-.017* (.01)	-.082** (.040)	-.18*** (.06)	.04 (.075)	-.11*** (.036)	-.11*** (.036)
%p.+65 in care home	.023*** (.007)	.086*** (.031)	.047 (.03)	.11** (.046)	.07** (.03)	.07** (.036)
Foreign nat.						
% French	-.013** (.005)	-.066** (.027)	.09 (.16)	-.08*** (.027)		
% German	-.001 (.004)	-.02* (.011)	-.22 (.33)	-.06*** (.018)		
% Dutch	.009* (.005)	.044* (.025)	.022 (.026)	.72*** (.22)		
% Italian	.028** (.011)	.028 (.053)	-.07 (.13)	.07 (.075)		
% Chinese	.047 (.17)	-.87 (.76)	1.06 (1.09)	-1.75** (.70)		
Regional D. Brus.=base						
Flemish	.003 (.10)	.21(.47)	-	-	.51 (.47)	.49 (.45)
Walloon	.18 (.11)	.98*(.51)	-	-	1.2** (.56)	1.1** (.54)
Constant	1.33** (.31)	1.12 (1.39)	4.47*** (1.45)	-2.12 (2.44)	1.57 (1.39)	1.78 (1.35)
R2	0.32	0.41	0.50	0.39		
N	581	581	300	262	581	581
Underid. test					13.2***	15.0***
Over ident. Sargan stat.						0.83 (0.36)
First stage					Cont.r. 31/3	Cont.r. 31/3
% Italian					-.08*** (.02)	-.075*** (.02)
% Chinese						.56* (.33)
F-test					12.9***	8.20***

Conclusions

- I employed linear as well as log-linear models. The latter capture the exponential growth of the contaminations better, but results are very similar to the linear model.
- The inclusion of the lagged dependent variable explains a lot of the variation of the dependent in my regression models. Hence it is important to find the correlates of the contamination rate early on in the epidemic.
- I do that by estimating the correlates of the start date of the epidemic as well as the correlates of the cumulative infection rate on March 31. I also turn my attention to the evolution of the epidemic by analysing the cumulative infection rate at the end of the strict lockdown (May 4)
- I find the same pattern. Income per capita, the share of elderly in the population, the share of elderly in home care and the exposure of the municipality to foreign travel, business and migration show up statistically significant in the analysis. Income in particular in the Flemish Region and foreign nationalities in particular in the Walloon Region.