INTERNET APPENDIX

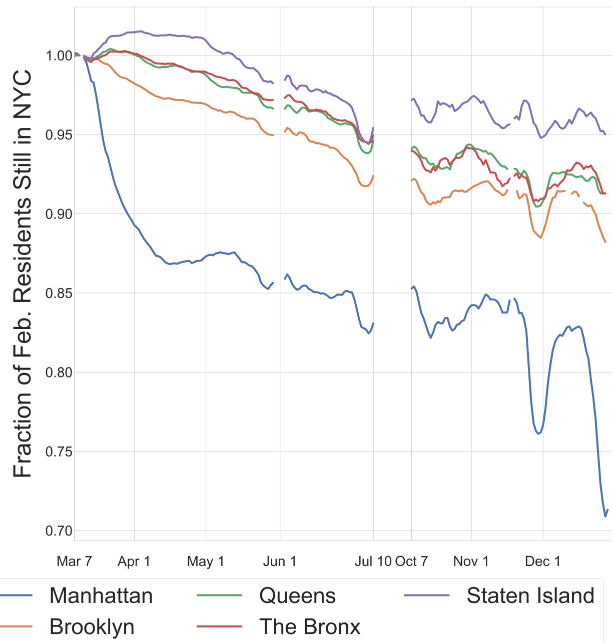
TABLE A1: Summary Statistics

	1st Quartile	Median	Mean	3rd Quartile
New Cases	2.49	5.69	23.51	14.90
New Cases per 1000 People	0.16	0.21	0.21	0.26
Net Inflow	49.49	103.72	223.34	222.72
Net Inflow per 1000 People	2.54	3.55	3.95	4.80
SCI-Instrumented Inflow	155.20	179.44	219.88	224.61
SCI-Instrumented Inflow per 1000 People	2.83	6.02	12.60	12.42
Social Connectedness Index	31.13	58.90	103.41	111.12
Social Connectedness Index (inflow-weighted)	70.67	182.44	793.69	542.70
Distance	312.91	428.23	505.35	594.83
Distance (inflow-weighted)	170.73	213.48	263.38	282.78
Number of nursing homes	2.00	5.00	7.65	8.00

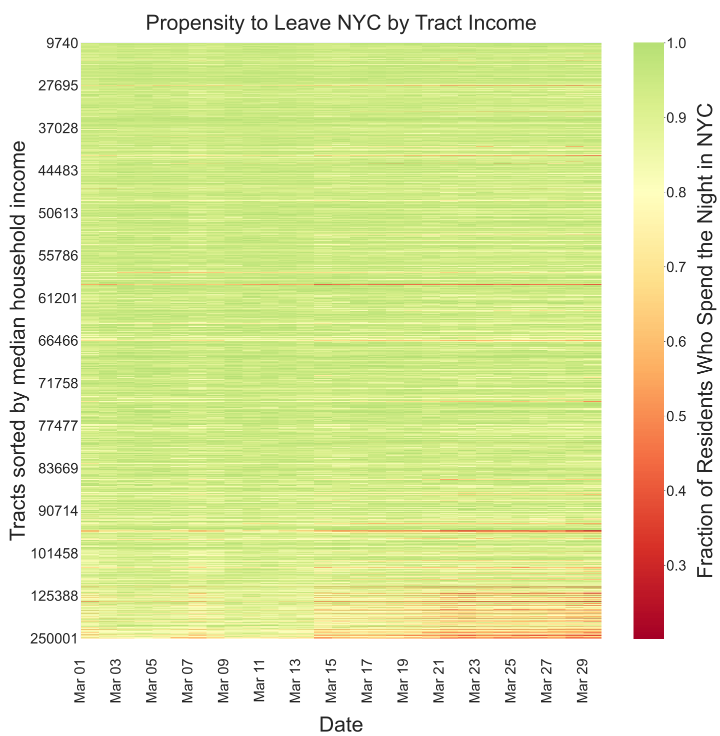
Summary statistics across all counties in our sample, with caseload and inflow data between March 1, 2020 through December 31, 2020. SCI and distance are calculated as the raw arithmetic average between the destination county and other counties from which there is a positive inflow. The inflow-weighted counterpart of these measures are calculated by using inflows from a particular county over total inflows as the weight in the weighted average.

FIGURE A1: Propensity to Leave NYC by Tract

Panel A: Propensity to Leave NYC by Borough (County)

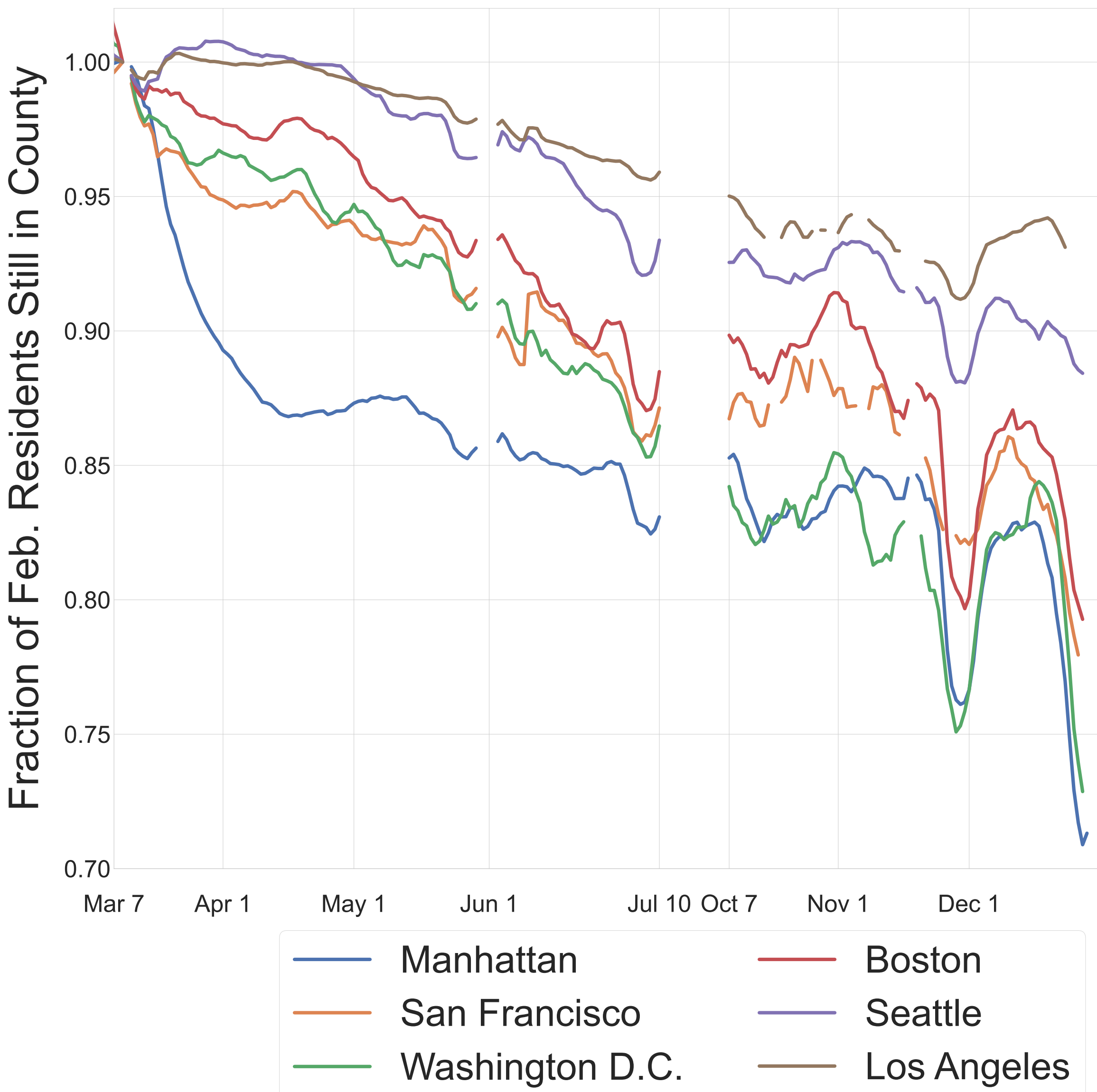


Panel B: Propensity to Leave NYC by Tract Income



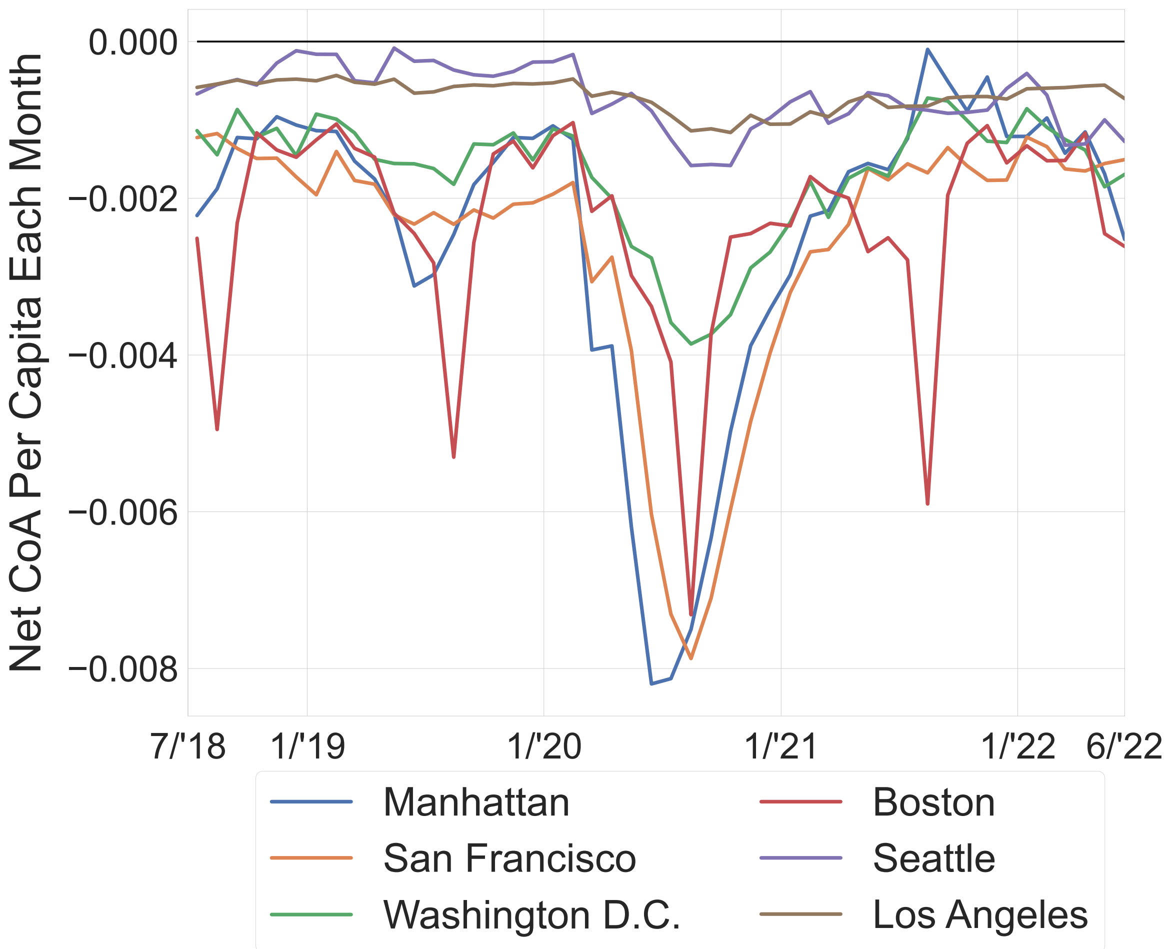
Notes: Panel A shows the propensity to leave NYC is a 7 day rolling average of the fraction of residents from a borough on a given date who spend the night outside of the five NYC counties divided by the total number of residents from that borough in the data that day, divided by the same fraction on March 9th, 2020. Panel B shows the same propensity but by tract, not normalized, and not a rolling average. Tract income data come from the Census ACS and are top coded at $250,001.

FIGURE A2: Propensity to Stay in City



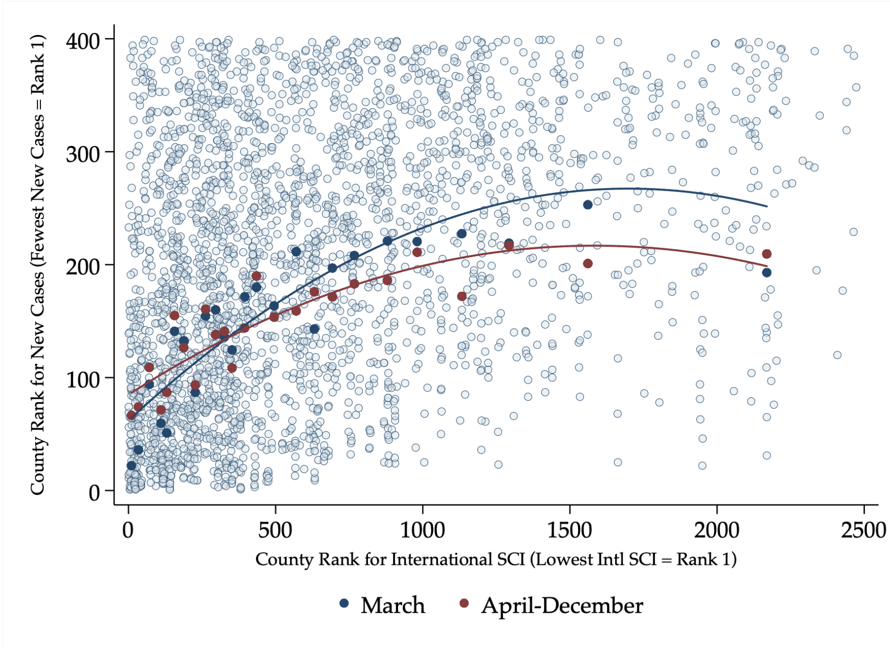
Notes: This figure shows the fraction of people who were identified as residents of their county in February, who were still in that county at later dates, when compared to those who are in the data at later dates but identified elsewhere in the country. Figure A1 Panel A shows the same for different boroughs of NYC.

FIGURE A3: Net Permanent Changes of Address Per Capita Each Month



Notes: This figure shows the net permanent changes of address each month from the USPS Change of Address data. Net is the number of people changing their address into a new zipcode minus the number changing their address out of the new zipcode. We aggregate this to the county level. "Permanent" is a USPS designation. Permanent moves tend to spike in the summer each year, lining up with the residential leasing calendar. Changes of address are scaled by county population in the 2018 1 year ACS. The y axis is not in percentage terms, so the largest spikes correspond to a 0.8% net permanent outflow that month.

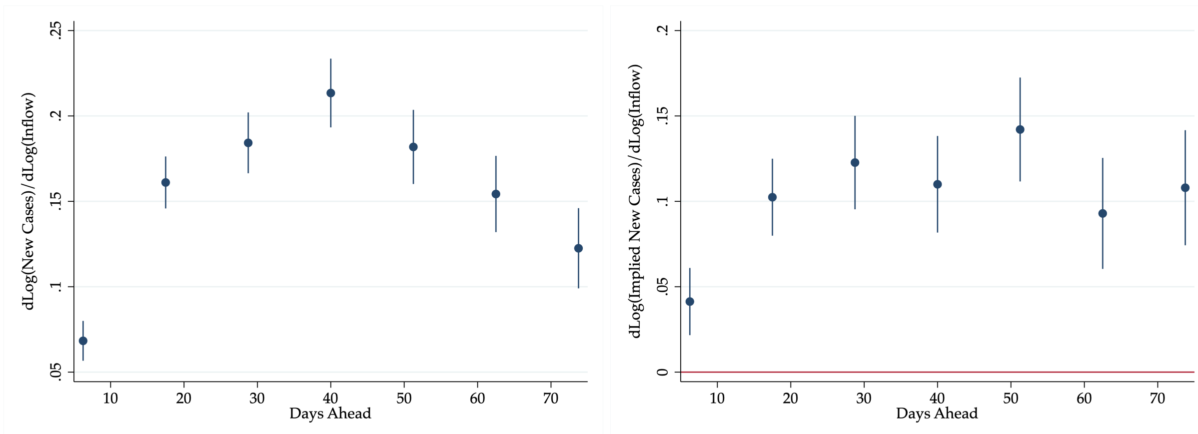
FIGURE A4: International Flows Seeded Initial US Cases



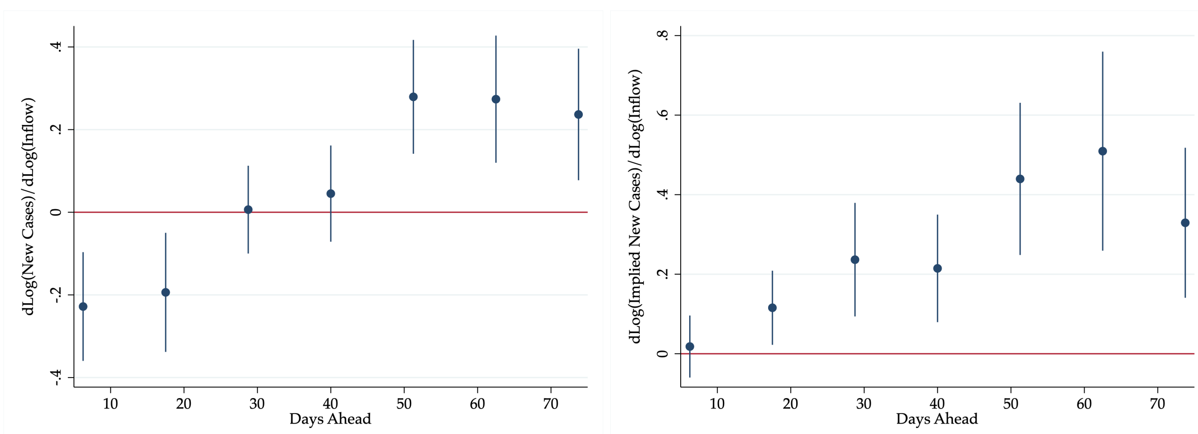
Notes: We rank counties by number of new cases each month and consider the top 10% of counties for each month from March to December. We also rank each county by the Facebook international SCI measure as a proxy for the likelihood that the county will receive inflows from abroad. We plot the international SCI rank on the x-axis and the new cases rank on the y-axis. The navy series represents the cross-sectional relationship between international SCI and new cases during March, and the maroon series represents the cross-sectional relationship between international SCI and new cases between April to December.

FIGURE A5: Impulse Response Functions: The Impact of Mobility on Log Case Growth

Panel A: OLS Results



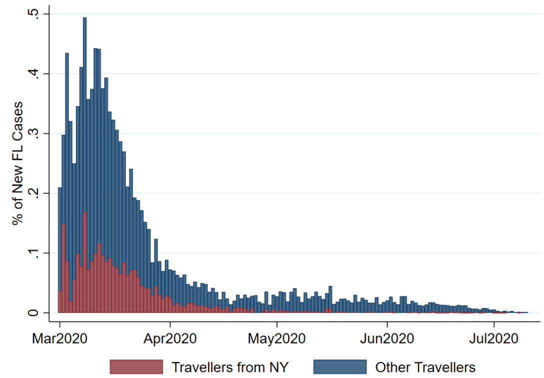
Panel B: IV Results



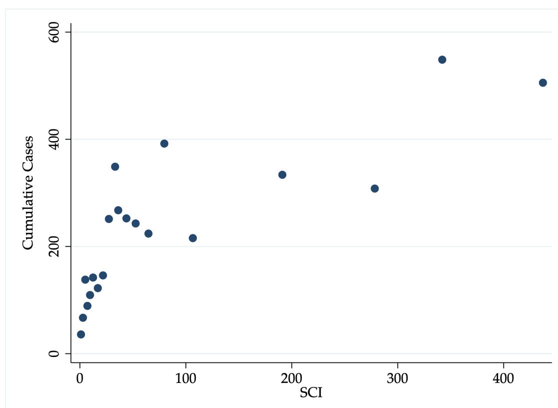
Notes: Impulse response functions with respect to inflow shocks. Following Wilson (2020), we run the following regression: $\log I_{i,t+h} - \log I_{i,t} = \beta_0 \cdot d\log I_{i,t} + \Sigma_{\tau=0}^{P}\beta_{h\tau} \cdot \text{Cumulative Cases}_{i,t-\tau} + \Sigma_{\tau=0}^{P}\rho_{h\tau} \cdot \log \text{Inflow}_{i,t} + \delta_i + \gamma_t + \epsilon_{i,t}$, where $I_{i,t}$ is the number of new cases in county i during time period t (10 days). We estimate this equation sequentially for horizons $h = 1$ to $h = 7$, and plot the sequence of coefficients $\beta_{h,0}$. The top panel shows the percentage increase in actual new cases (left) and new cases implied from deaths (right) after a 1% increase in inflow. Implied new cases from deaths on date t are equal to deaths on date t + 21. This follows Wilson (2020) to deal with potential under reporting and mistiming of cases. Panel B shows the same outcome variables after a 1% increase in SCI-instrumented inflows. For each period (10 days), we control for the cumulative case load at that point in time, and include county FEs and period FEs.

FIGURE A6: The Impact of Inflows on New Cases in Florida

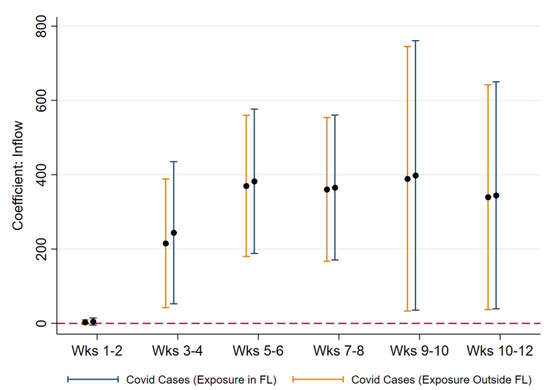
Panel A: Many Cases in March Were Associated With Recent Travel



Panel B: People Testing Positive After Travel Came From Regions With High SCI to FL



Panel C: Inflows Predict Cases Even When Removing Those Who May Have Gotten Exposed While Traveling, Suggesting Spillovers Due To Travel



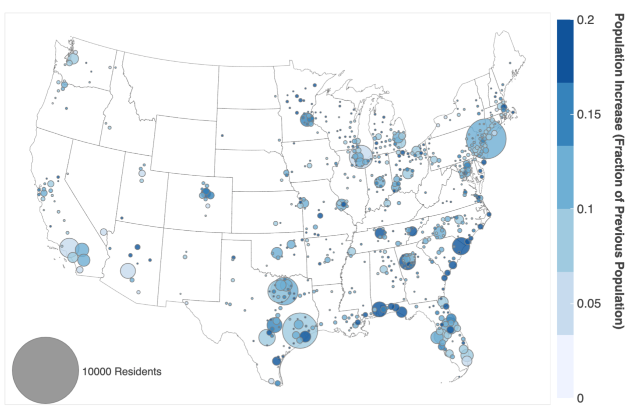
Notes: Panel A shows the proportion of all reported cases which where the patient had a known history of exposure to covid associated with travel outside of Florida. In Panel B, we plot a binscatter of the SCI between the Florida county and the originating state of a traveller with a covid case and the cumulative travel-related cases seen in the county by July 2020. Panel C shows the coefficient of SCI-instrumented inflows on total cases (both travel and non-travel) and only non-travel-related cases. We limit our sample to the period of time where travel-related cases account for at least 5% of all reported cases and show the coefficient for successive two-week periods from the beginning of March to the end of May.

FIGURE A7: Nationwide Inflow at the County Level

Panel A: March

Panel B: April

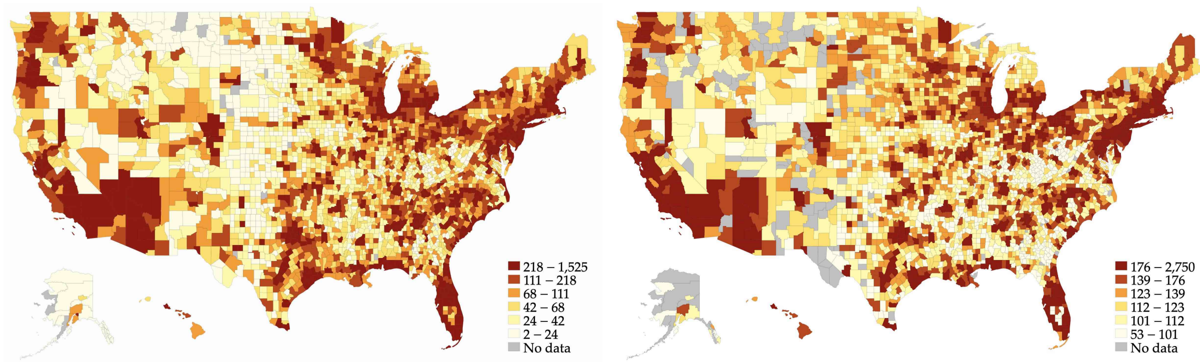
Panel C: May



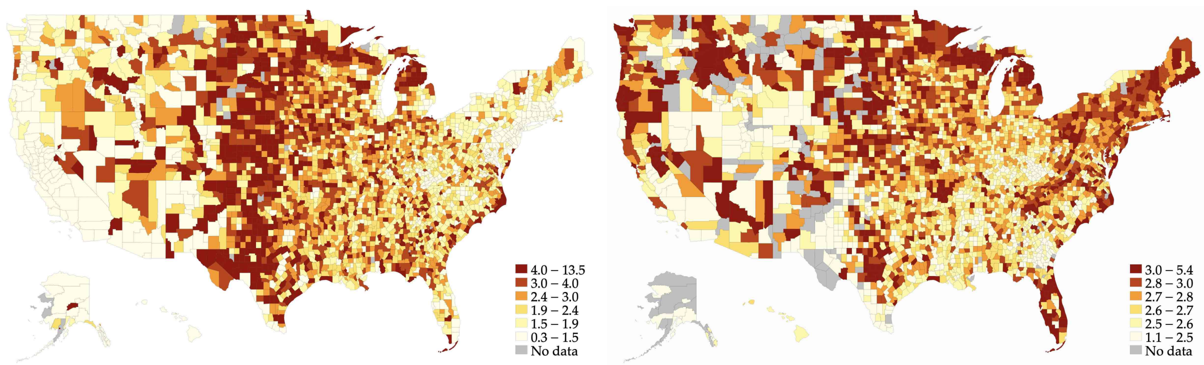
Notes: A person counts as outflow if they were resident in a county in the previous month, and then on a given day in the current month, have a modal night county somewhere else. We count the number of people who were residents of a county in the previous month, who are present in the data on a given date in the current month, who are in a different county. We average this across the last week of the month to get our outflow number. The population change fraction is this number divided by the total number of residents of that county who were in the data on that date, also averaged across the last week of data. Inflow counts the number of people who were not residents in a county in the previous month, who have the county is their modal night county on a given date. It is averaged the same way outflow is. Net flow is inflow minus outflow.

FIGURE A8: Raw vs. SCI-Fitted Inflows

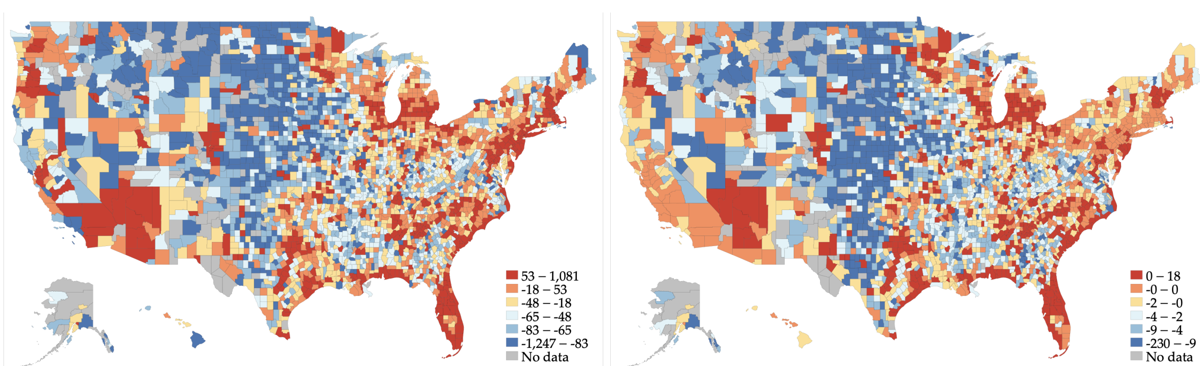
Panel A: Total Inflow



Panel B: Inflow Per Capita

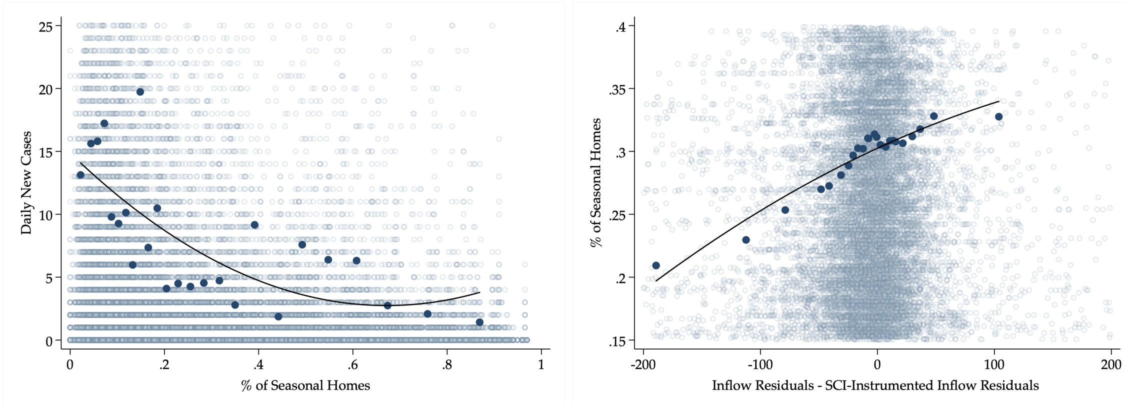


Panel C: Difference Between Raw and SCI-Instrumented Inflows



Notes: Panels A & B show raw inflows and raw per capita inflows (left column) and SCI-fitted inflows and inflows per capita (right column). Darker red represents higher values; each map is colored by splitting the data into sextiles, resulting in different cutoff values for different colors, which vary across maps. Panel C shows the difference between raw and SCI-instrumented inflows (inflow − SCI-instrumented inflows). The left side of Panel C shows the difference in levels, and the right side shows the difference in per capita terms. Red counties represents regions where raw inflow is higher than instrumented inflow, and blue counties represent regions where raw inflow is lower than instrumented inflow.

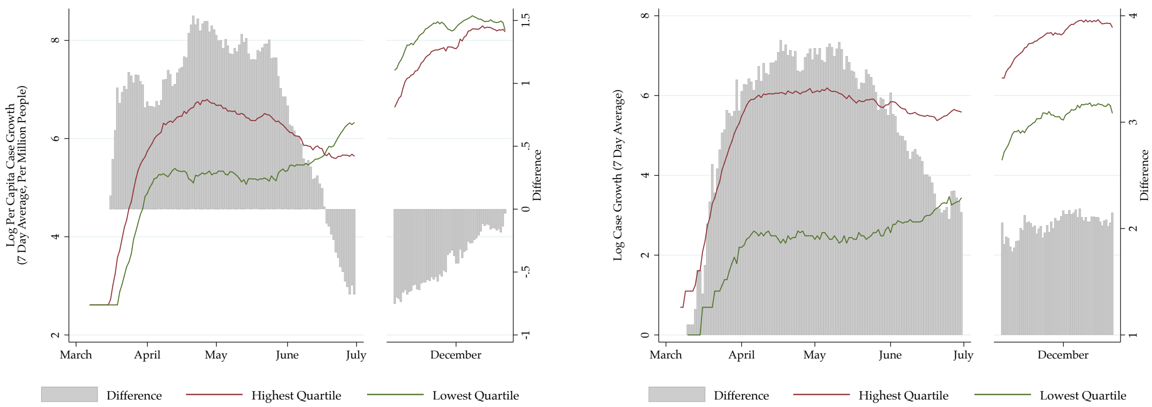
FIGURE A9: Differences in IV and OLS Estimates Due to Differences in Instrumented and Raw Inflow



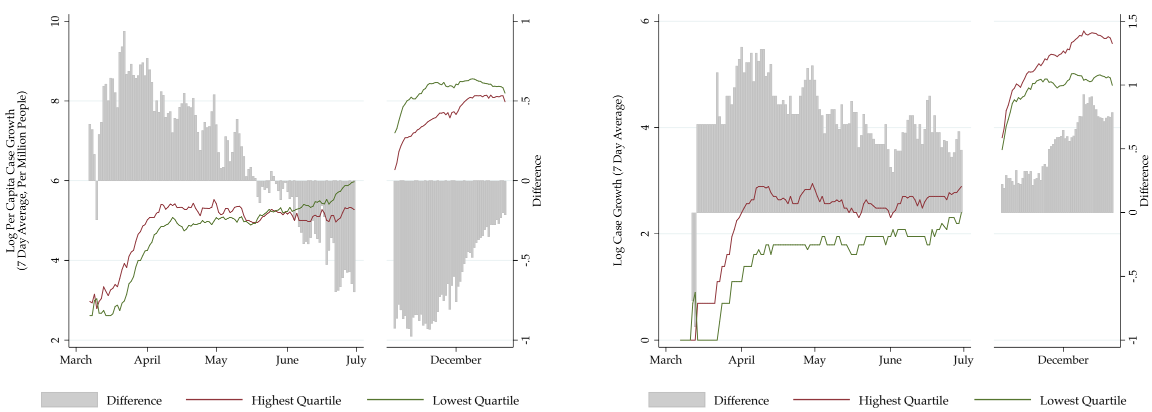
Notes: The left binscatter shows new cases tend to be lower in regions with a greater proportion of seasonal homes. The right binscatter shows the unexplained portion of raw inflows (after partialling out the impact of the control variables we use in our baseline specification) tends to be higher than the unexplained portion of instrumented inflows, precisely in regions with more seasonal homes, and hence, regions with lower new cases. We point to seasonal homes as an imperfect measure of a possible omitted variable, which is positively correlated with the raw (not instrumented) inflow and negatively correlated with new cases.

FIGURE A10: March Inflows From NYC vs. Log Case Growth

Panel A: Large Central, Fringe, and Medium Metropolitan Areas



Panel B: Micropolitan and Non-Core Areas



Notes: Counties are split into quartiles based on total inflows from New York City during the month of March. The two charts on the left show median log(total cases per capita growth) over time for counties in the fourth quartile and the first quartile of inflows. The two charts on the right show log(total case growth). Urban classification based on the NCHS urban-rural classification scheme: large central metros, large fringe and medium metropolitan areas (categories 1, 2, & 3) and micropolitan and non-core areas (categories 4, 5 & 6). This figure differs from IV by considering case growth, rather than totals.

TABLE A2: Impact of Migration on New COVID-19 Cases: Standardized Coefficients

Panel A: New COVID-19 Cases Against Inflow

	OLS						IV					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Inflow (standardized)	0.240***	0.015	-0.020				1.183***	0.378***	-0.098			
	(0.011)	(0.012)	(0.013)				(0.061)	(0.048)	(0.128)			
Per Capita Inflow (standardized)				0.001	0.002	0.002				5.650***	0.574***	1.090***
				(0.001)	(0.001)	(0.002)				(0.510)	(0.144)	(0.155)
High Incoming Cases	0.080***	0.083***	0.083***				0.093***	0.088***	0.085***			
	(0.002)	(0.002)	(0.002)				(0.003)	(0.002)	(0.003)			
High Incoming Cases Per Capita				0.513***	0.519***	0.519***				0.662***	0.533***	0.546***
				(0.003)	(0.003)	(0.003)				(0.020)	(0.005)	(0.006)
Far Indicator	0.166***	0.104***	0.065***	0.030***	0.019*	0.019*	0.185***	0.126***	-0.111	0.013	0.061***	-0.041*
	(0.015)	(0.013)	(0.012)	(0.010)	(0.010)	(0.010)	(0.018)	(0.013)	(0.118)	(0.064)	(0.016)	(0.022)
Far Indicator × Inflow (standardized)			0.253***						1.455*			
			(0.023)						(0.782)			
Far Indicator × Per Capita Inflow (standardized)						0.002						-0.520***
						(0.005)						(0.050)
All Controls	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y
State x Month x Population Decile FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	595,945	573,968	573,968	595,945	573,968	573,968	595,945	573,968	573,003	595,945	573,968	573,968
F							5,088	2,738	1,490	150	297	171

Columns 1-3 show our main regression specification 1, where all variables (excluding indicator variables) are standardized to have zero mean and unit variance. Columns 4-6 repeat the exercise using inflow per capita as the explanatory variable. Columns 7-12 repeat the exercise for columns 1-6, where inflow is instrumented with the weighted SCI, as in 2. The sample period is March 1, 2020 through December 31, 2020. Standard errors are in parentheses, and * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE A3: March Inflows from NYC vs. Log New Cases

	Panel A: Log(New Cases) Against Inflow from NYC			
	OLS		IV	
	(1)	(2)	(3)	(4)
High NYC Inflow	0.702***	0.128***	6.379***	2.956***
	(0.012)	(0.008)	(0.221)	(0.122)
Controls	Y	Y	Y	Y
State × Month × Population Decile FE	Y	Y	Y	Y
N	100,595	99,846	100,595	99,846
	Panel B: First Stage Estimates			
Weighted SCI			0.011	0.159
			(0.0003)	(0.0055)
F			1,364	334

This table shows our baseline regression, where the outcome is the log(New Cases) and the inflow variable is replaced with an indicator equal to 1 if a county was in the top quartile of all counties which received inflows from NYC in March. The sample period is March 1, 2020 through December 31, 2020. Standard errors are in parentheses, and * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance.

TABLE A4: Total Inflows from NYC vs. New Cases

Panel A: New COVID-19 Cases Against Inflow from New York City

	OLS						IV					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Inflow from NYC	157.542***	71.257***	67.542***				2400.411***	410.191***	209.346***			
	(18.373)	(11.464)	(11.164)				(133.485)	(58.416)	(61.673)			
Per Capita Inflow From NYC				16687.929	-50654.054**	-58041.353**				13252.582***	1799.347***	-547.578
				(25829.686)	(25408.079)	(25020.414)				(607.568)	(408.289)	(390.184)
High Incoming Cases	11040.275***	12273.657***	12197.325***				3871.768***	11136.293***	11250.140***			
	(256.317)	(278.397)	(280.875)				(836.617)	(339.744)	(380.709)			
High Incoming Cases Per Capita				209.472***	211.994***	211.998***				0.196***	0.210***	0.212***
				(1.281)	(1.293)	(1.293)				(0.002)	(0.001)	(0.001)
Far Indicator	22711.540***	14836.297***	9144.736***	11.804***	7.575*	7.222*	18154.358***	14264.012***	-21338.148***	-0.021***	0.005	-0.114***
	(2337.384)	(1992.346)	(2060.395)	(3.920)	(4.049)	(4.090)	(2332.577)	(2001.609)	(2968.682)	(0.005)	(0.004)	(0.013)
Far Indicator × Inflow from NYC			3953.143***						24935.979***			
			(617.296)						(2476.870)			
Far Indicator × Per Capita Inflow From NYC						66887.757						22423.299***
						(1.05e+05)						(2217.196)
All Controls	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y
State x Month x Population Decile FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	595,945	573,968	573,968	595,945	573,968	573,968	595,945	573,968	573,968	595,945	573,968	573,968

Panel B: First Stage Estimates of Inflow on SCI

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Weighted SCI							10088.338***	12600.925***	13152.353***	0.001***	0.001***	0.001***
							(346.688)	(460.555)	(477.527)	(0.00002)	(0.00002)	(0.00002)
F							881	748	395	1,418	1,049	523

Columns 1-3 show our main regression specification 1, with inflows from the five NYC boroughs as the explanatory variable rather than total inflows. Columns 4-5 repeat the exercise using inflow per capita as the explanatory variable. Columns 7-12 repeat the exercise for columns 1-6, where inflow from New York City is instrumented with the weighted SCI, as in 2. The sample period is March 1, 2020 through December 31, 2020. Standard errors are in parentheses, and * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance. Note that all coefficients and standard errors in Panel A are scaled up by 1×10^3, with the exception of columns 4-6, where coefficients and standard errors are scaled up by 1×10^6. All coefficients and standard errors in Panel B are scaled up by 1×10^6.

TABLE A5: Impact of Migration on New COVID-19 Cases with Clustering at the Commuting Zone Level

Panel A: New COVID-19 Cases Against Inflow

	OLS						IV					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Inflow	98.388***	6.142	-8.367				484.466***	154.914***	-40.224			
	(29.551)	(5.433)	(7.340)				(147.766)	(54.053)	(72.277)			
Per Capita Inflow				177.069	255.225	233.067				750.188	76.223	144.744
				(357.738)	(338.459)	(352.469)				(579.412)	(88.688)	(107.333)
High Incoming Cases	11975.997***	12314.294***	12320.634***				13933.822***	13166.577***	12640.649***			
	(1297.772)	(1253.155)	(1253.578)				(1894.675)	(1395.443)	(1525.762)			
High Incoming Cases Per Capita				209.552***	211.982***	211.981***				0.271***	0.218***	0.223***
				(5.044)	(5.130)	(5.130)				(0.050)	(0.009)	(0.010)
Far Indicator	24729.038*	15438.162***	-12824.185**	12.105*	7.697	6.800	27584.627*	18805.957***	-1.461e+05*	0.005	0.025	0.261***
	(13674.285)	(5376.889)	(5095.142)	(6.260)	(5.043)	(6.759)	(15566.115)	(6542.377)	(87207.827)	(0.088)	(0.023)	(0.093)
Far Indicator=1 × Inflow			103.759***						595.945*			
			(29.132)						(307.908)			
Far Indicator=1 × Per Capita Inflow						278.969						-69.070***
						(1188.784)						(24.568)
All Controls	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y
State x Month x Population Decile FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	595,945	573,968	573,968	595,945	573,968	573,968	595,945	573,968	573,003	595,945	573,968	573,968
F							99	74	49	1	5	2

Panel B: First Stage Estimates of Inflow on SCI

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Weighted SCI							42277.757***	32742.431***	34931.996***	0.009	0.015**	0.013*
							(4243.984)	(3801.519)	(3691.353)	(0.008)	(0.007)	(0.007)
F							112	85	74	0	2	3

Columns 1-3 shows our main regression specification1. Columns 4-6 repeat the exercise using inflow per capita as the explanatory variable. Columns 7-12 repeat the exercise for columns 1-6, where inflow is instrumented with the weighted SCI, as in 2. The sample period is March 1, 2020 through December 31, 2020. Standard errors are in parentheses, and * denotes 10% significance, ** denotes 5% significance, *** denotes 1% significance. Note that all coefficients and standard errors in Panel A are scaled up by 1×10^3, with the exception of columns 4-6, where coefficients and standard errors are scaled up by 1×10^6. All coefficients and standard errors in Panel B are scaled up by 1×10^6. Estimates in this table have standard errors clustered at the Commuting Zone level.