Urban Economics: Introduction

Matthew Turner

Brown University

February 4, 2020

Objectives

This course is about two main topics:

- How can we use economics to understand the internal structure of cities. For example, why are they denser and more expensive at their centers? Why has manufacturing migrated from the center to the edge over the past 100 years?
- How can we use economics to understand the size of cities, their locations, and the distribution of activities across them?

These questions are important for three reasons

- Cities are important: Many people live in them.
- Urbanization is related to development and growth.
- The tools we develop will help us to evaluate the implications of policies with a ‘spatial’ component; school choice, development zones, MTO and transportation infrastructure.

Many people live in them/Development and growth

	World population	Urban Population	Urb. Share
1960	3b	1b	0.34
2012	7b	3.5b	0.52

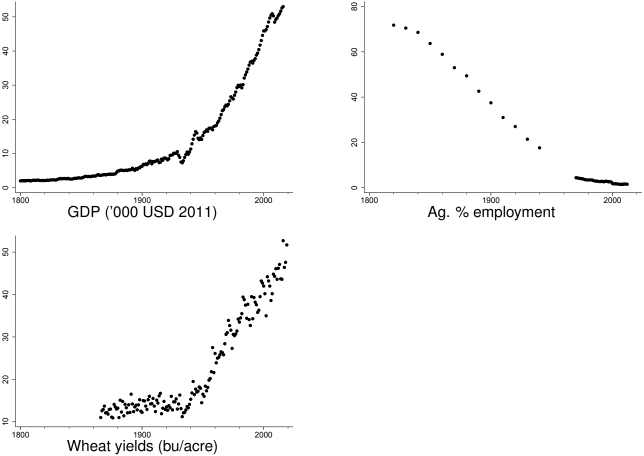
- There are more people in cities today than there were people in the world in 1960.
- 80% of world economic activity is in cities. With only half the people in cities, this means an average urban resident in about 4 times as productive as an average rural resident.

If you are interested in growth or development, you should probably be interested in cities.

Spatial equilibrium is useful

- Chetty and Hendren (2016) and Kling, Liebman and Katz(2007) investigate the effect of a randomly assigned subsidy (the MTO experiment) that encourages poor households to move to better neighborhoods. How do we think about the welfare implications of such a policy if we do not change the number of housing units?
- Which children live in houses with lead paint? What are the implications of an expensive remediation mandate for rental housing for the level and distribution of exposure?
- ‘Opportunity zones’ provide tax cuts for capital investments in ‘poor’ census tracts. Do these zones create opportunity? Do they create opportunity for the intended population? Do they shift employment from one place to another?

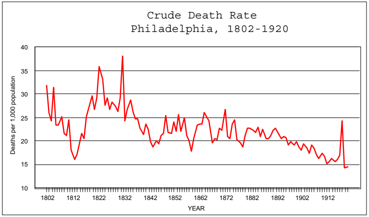
Urbanization in the US, basic facts

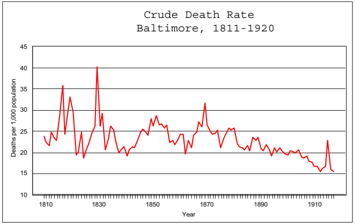


(a)Maddison Project GDP data, (b)Historical Statistics of the United States, 1789-1945, Ch D, (c) of Agriculture (2020).

In 1860, the US was about 40% urban (100%-60% Ag.) GDP was about 3400 USD2011. In 1880 the US was 50% urban with GDP about 4900 USD2011.

Urbanization in the US





(a)Haines (2001)

From Haines (2001), for every rural person in 1000 who died, about 1.25 died in a city. By 1940, this number had reached parity. Current US crude death rate $\sim 8.5/1000$.
`https://data.worldbank.org/indicator/SP.DYN.CDRT.IN?locations=US` (Jan 27, 2020)

Urbanization in the US

Story

- At low levels of income, urban population is limited by agricultural surplus. This may have been a factor in the 18th and 19th century, probably not after 1950.
- Basic story: Farmers migrate from the countryside to progressively more productive factory jobs in the cities.
- The high death rate in cities slows this process.

Spatial model

We just described a model of spatial equilibrium. Two locations, U and R, indexed by j. A continuum of people, indexed by i.

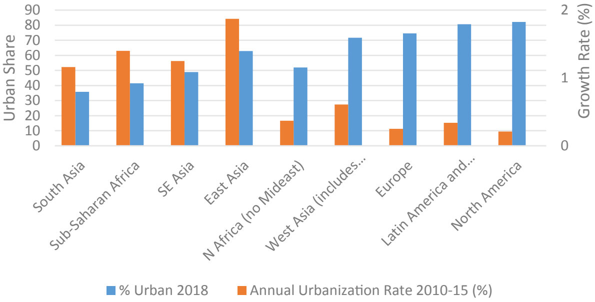
- Income for i at j is I_{ij}.
- Amenity at j, A_j, a location specific public good (e.g., death rate).
- Individual specific 'affinities' for U and R, $(\epsilon_{iU}, \epsilon_{iR})$.
- Cost to migrate between locations, τ.

Spatial equilibrium is an arrangement of people such that,

$$\begin{aligned} V(I_{iU}, A_U, \epsilon_{iU}) &\geq V(I_{iR} - \tau, A_R, \epsilon_{iR}) && \forall i \in U \\ V(I_{iR}, A_R, \epsilon_{iR}) &\geq V(I_{iU} - \tau, A_U, \epsilon_{iU}) && \forall i \in R \end{aligned}$$

This is, more-or-less, the Rosen-Roback model, and is one of the more important ideas in urban economics. Note the fundamentals: Income, Amenities, travel costs, individual heterogeneity. What happens as $\tau \longrightarrow \infty$? As $var(\epsilon) \longrightarrow \infty$?

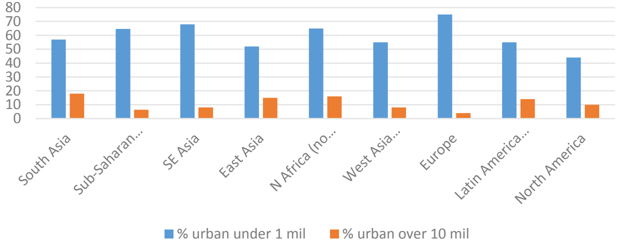
Urbanization in the developing world, basic facts



Regions are UN regions. The Middle East is part of West Asia (not North Africa) and Latin America includes the Caribbean. Oceania is excluded. Based on Henderson and Turner (2020).

We are building cities in Asia and Africa. Everywhere else, urbanization seems about done.

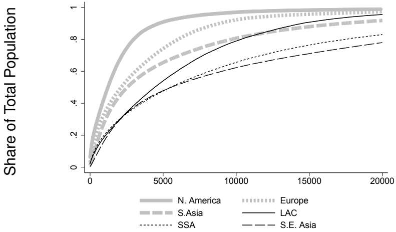
Urbanization in the developing world

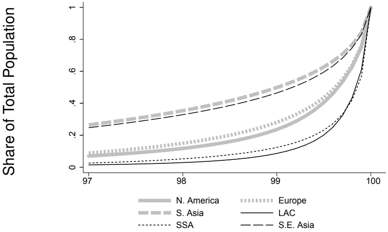


Regions are UN regions. The Middle East is part of West Asia (not North Africa) and Latin America includes the Caribbean. Oceania is excluded. Based on Henderson and Turner (2020).

The size distribution of cities is different across regions. Why are really big cities more important in Asia and North Africa?

Population Density and Land Use by Region

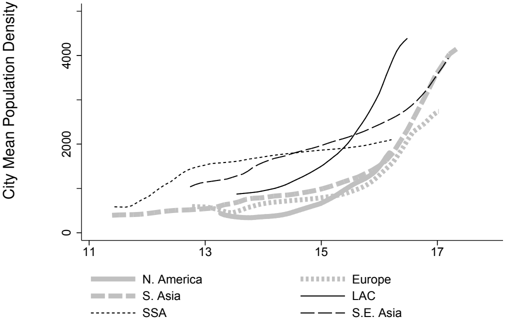




(left) Cumulative share of population by density. (right) Cumulative share of population by land area in the region. Based on population data from GHS. Based on Henderson and Turner (2020)

Population densities are much higher in poorer countries. Small shares of land are occupied everywhere. Caveat: GHS data is suspicious.

City mean population density



Vertical axis is mean population density from GHS in a 50km radius disk centered on the centroid of each of the 657 UN world cities. Horizontal axis is total population in the same disk, also from GHS. Based on Henderson and Turner (2020)

The density of African cities does not increase with size. Everywhere else it does. Same caveat, this is from GHS.

Share of manufacturing in GDP by region and year.

Region	1990	2000	2010	2017
E. Asia	24.6	25.2	27.6	27.4
S.E. Asia	22	24.8	22.6	20.9
L. America and Caribbean	20.7	17.9	15.7	15.2
N. Africa	17.6	17.9	16	16
Europe	17.5	15.3	11.9	11.8
S. Asia	15.9	15.6	16.1	14.4
W. Asia	14.4	13.2	12.1	13.8
S.S.A.	13.8	11.6	8	9

From Henderson and Turner (2020). Data from the World Development Indicators 2018 are organized by UN regions. The table reports regional weighted averages using weights based on country share of regional GDP in 2017. Data cover 126 countries in a consistent sample over time. The Middle East is part of West Asia (not North Africa). Oceania is excluded

Farmers in African cities by city size.

African countries	All urban	All rural	Primate city	Secondary cities (top 25%)	Tertiary cities (50-75%)	All others
% reporting agriculture as main industry	20.5	88	8.5	23.8	38.6	41.3
% reporting manufacturing as main industry	10.6	<2	12.4	10	8.3	7.3

From Henderson and Turner (2020). The data are from IPUMS for the most recent census for Ethiopia, Tanzania, Uganda, Mozambique, Ghana, Cameroon, Mali, Malawi, Zambia, Sierra Leone, Liberia, and Botswana. Small cities are in the bottom 50% of cities by size and tertiary cities are in the 50-75th percentiles. Cities are defined by night-light boundaries to which population is assigned. Details are reported in Henderson and Kriticos (2018).

Urban Share and national income

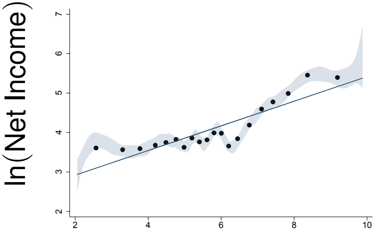
- SSA: 40% urban share $\sim$ 2010, GDPpp$\sim$ 1481 USD1990.
- Latin America: 40% urban share $\sim$ 2010, GDPpp$\sim$ 2500 USD1990.
- East Asia: 60% urban share $\sim$ 2000, GDPpp$\sim$ 5451 USD1990.
- United States: ag. employment $\sim$ 1860, GDPpp$\sim$ 3400 USD2010.

Latin America and SSA are building cities when they are much poorer than were the US and Europe when they were at similar shares of urban population Henderson and Turner (2020).

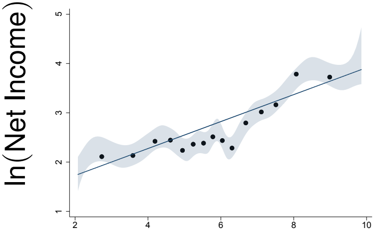
Are cities urbanizing too early?

- Story: on the basis of developed world experience, we think farmers move to cities to take high paying manufacturing jobs but are subject to high rates of disease.
- This seems not to apply in SSA. Manufacturing is scarce, farmers live in cities, but cities are probably too poor to manage 'demons of density'(Glaeser, 2011).
- But cities are growing fastest in SSA!
- Cities in SSA and Asia are different from those in the US and Europe. Are different economic forces at work? Are SSA cities growing 'too fast' to rationalize with spatial equilibrium?
- ... or maybe, we don't have our facts straight.

Log net income versus log population density/km^2 within a 5k radius.



(a) No Controls

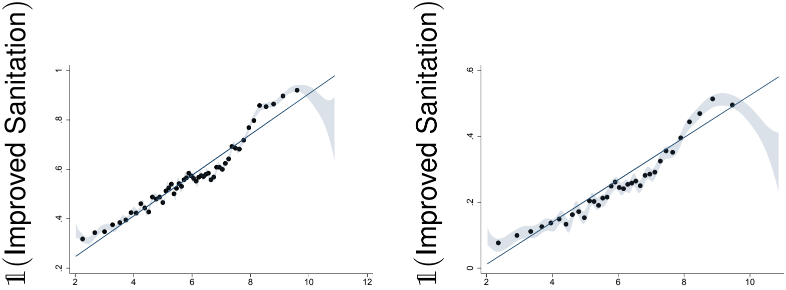


(b) Demographic controls

Binscatter plots of LSMS net income of respondent household against the log of GHS population density in a 5km disk around the survey respondent. Log population density is censored below 2. Left panel has no controls. Right panel includes demographic controls and country fixed effects. Shading indicates 95% confidence band. Income includes wage income, net farm income and net business income. For a small number of observations expenses exceed (monthly) incomes. We drop these observations to permit logarithmic scaling.

Income increases dramatically with density.

Access to improved sanitation versus log population density/km^2 within a 5k radius.



(a) No Controls

(b) Demographic controls

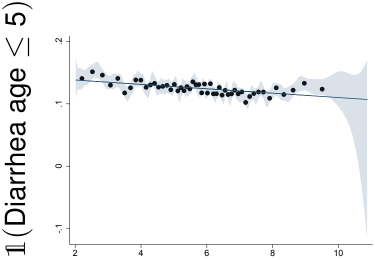
Binscatter plots of a DHS indicator variable that is one if a respondent household has access to improved sanitation. Log population density is censored below 2. Left panel is unconditional. Right panel includes demographic controls and country fixed effects. Shading indicates 95% confidence band.

Access to improved sanitation increases rapidly with city size. NB: ‘improved sanitation’$\neq$ ‘flush toilets’.

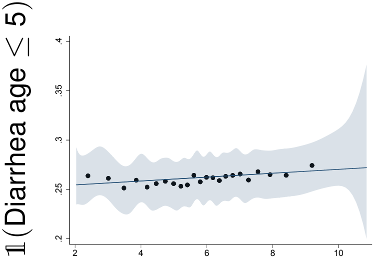
Are cities urbanizing too late?

- Wages are much higher in SSA cities. Moving from 550 people per km (ln $\sim$ 6.3) to 8100 (ln $\sim$ 9) increases net income by about a factor of 4. Whatever people are doing in cities, they are much more productive than they are in the countryside.
- Access to improved sanitation increases very rapidly with density. In spite of their poverty, SSA cities are providing basic public services to most of their residents.
- Wages are better, public services are better, why don't more people move?
- Can we rationalize this 'too slow' urbanization (Gollin, Kirchberger, Lagakos (2017)) with spatial equilibrium? Maybe people are really attached to their homes or migration is really expensive?
- ... or maybe, we don't have our facts straight.

Diarrhea last two weeks for children ≤ 5 vs log pop. density/km^2 within a 5k radius.



(a) No Controls



(b) Demographic controls

Binscatter plots of a DHS indicator that is one if a child five or under had diarrhea in the past two weeks against the log of GHS population density in a 5km disk around the survey respondent. Log population density is censored below 2. Left panel is unconditional. Right panel includes demographic controls and country fixed effects. Shading indicates 95% confidence band.

Rate of illness in children increases with density, ceteris paribus.

Density gradients for Afrobarometer, LSMS and DHS outcomes.

	No controls		Controls						
Outcome	β *s.e.*	R^2	β *s.e.*	R^2	$\overline{y}$ *s.e.*	$\overline{x}$ *s.e.*	N	Clusters	Countries
Data: *LSMS*									
ln(Income)	.3126[a]	0.067	.3170[a]	0.856	4.097	5.77	35,231	2,118	5
	(.0161)		(.0141)		(2.014)	(1.67)			
ln(Wage)	.1177[a]	0.019	.0488[a]	0.553	1.191	6.38	18,806	1,704	5
	(.0152)		(.0094)		(1.435)	(1.69)			
Controls: $\mathbb{1}$(*Kindergarten*), $\mathbb{1}$(*Some prim. sch.*), $\mathbb{1}$(*Some high sch.*), *age O(2)*, $\mathbb{1}$(*fem.*).									
Data: *DHS household*									
Electricity	.0797[a]	0.084	.0444[a]	0.827	.691	5.96	987,081	28,088	38
	(.0012)		(.0010)		(.462)	(1.68)			
Safe Water	.0853[a]	0.083	.0576[a]	0.655	.510	5.95	1,005,468	28,604	39
	(.0013)		(.001)		(.500)	(1.69)			
Imp. Sanitation	.0825[a]	0.079	.0630[a]	0.662	.572	5.95	1,005,283	28,604	39
	(.0010)		(.0010)		(.495)	(1.69)			
Controls: *H.H. size O(2)*, $\mathbb{1}$(*fem. HoH*), *age HoH O(2)*, $\mathbb{1}$(*Some prim. sch. HoH*), $\mathbb{1}$(*Some sec. sch. HoH*), $\mathbb{1}$($>$ *sec. sch. HoH*).									
Data: *DHS school*									
School$\geq$8yr	.0497[a]	0.029	.0158[a]	0.719	.611	5.94	95,687	25,529	39
	(.0014)		(.0011)		(.488)	(1.67)			
Controls: $\mathbb{1}$(*fem.*), $\mathbb{1}$(*fem. HoH*), *age HoH O(2)*, $\mathbb{1}$(*Some prim. sch. HoH*), $\mathbb{1}$(*Some sec. sch. HoH*), $\mathbb{1}$($>$ *sec. sch. HoH*).									
Data: *DHS female*									
Contraception	.0297[a]	0.011	.0122[a]	0.595	.496	5.9	183,273	19,294	36
	(.0016)		(.0009)		(.500)	(1.76)			
Justified Beating	-.0361[a]	0.017	-.0120[a]	0.499	.384	5.87	575,495	20,129	39
	(.0016)		(.0009)		(.486)	(1.76)			
Victim	.0001	0.000	.0074[a]	0.320	.277	5.8	194,157	17,951	31
	(.0010)		(.0009)		(.448)	(1.77)			
Tot. # births	-.0278[a]	0.008	-.0109[a]	0.370	.298	6.01	1,110,331	28,604	39
	(.0007)		(.0004)		(.531)	(1.68)			
Controls: *age O(2)*, $\mathbb{1}$(*Some prim. sch.*), $\mathbb{1}$(*Some sec. sch.*), $\mathbb{1}$($>$ *sec. sch.*), $\mathbb{1}$(*fem. HoH*), *age HoH O(2)*, $\mathbb{1}$(*Some prim. sch. HoH*), $\mathbb{1}$(*Some sec. sch. HoH*), $\mathbb{1}$($>$ *sec. sch. HoH*).									
Data: *DHS birth*									
Infant Death	-.0006[a]	0.000	.0008[a]	0.038	.035	5.75	294,385	28,205	39
	(.0002)		(.0002)		(.184)	(1.71)			
Controls: $\mathbb{1}$(*fem.*), *age (mother) O(2)*, $\mathbb{1}$(*Some prim. sch.(mother)*), $\mathbb{1}$(*Some sec. sch.(mother)*), $\mathbb{1}$($>$ *sec. sch.(mother)*), $\mathbb{1}$(*fem. HoH*), *age HoH O(2)*, $\mathbb{1}$(*Some prim. sch. HoH*), $\mathbb{1}$(*Some sec. sch. HoH*), $\mathbb{1}$($>$ *sec. sch. HoH*).									

Note: *Regressions of respondent level 'outcome' on log population density in a 5km disk. Standard errors are clustered by 'survey cluster'. Each row reports results from two regressions, one without demographic controls and one with;* $^a = 1\%$, $^b = 5\%$, $^c = 10\%$, *all two-tailed tests. Relevant demographic controls are listed at the bottom of each panel.* $\overline{y}$ *and* $\overline{x}$ *are mean of outcome and* ln(*pop. density*) *in the 'no-controls' sample. Except for the LSMS panel, we lose only a tiny number of observations when we add controls.(Henderson and Turner, 2020)*

Density gradients for Afrobarometer, LSMS and DHS outcomes.

	No controls		Controls						
Outcome	β s.e.	R^2	β s.e.	R^2	$\overline{y}$ s.e.	$\overline{x}$ s.e.	N	Clusters	Countries
Data: *DHS children*									
Diarrhea	-.0035[a]	0.000	.0030[a]	0.160	.125	5.76	512,855	28,507	39
	(.0005)		(.0004)		(.331)	(1.71)			
DPT3	.0209[a]	0.007	.0123[a]	0.798	.763	5.76	95,334	24,914	39
	(.0013)		(.0011)		(.425)	(1.71)			
Cough	-.0001	0.000	.0038[a]	0.255	.188	5.76	513,082	28,507	39
	(.0008)		(.0006)		(.391)	(1.71)			
Controls: *age O(2),* $\mathbb{1}($*Some prim. sch.(mother)*$)$, $\mathbb{1}($*Some sec. sch.(mother)*$)$, $\mathbb{1}(>$ *sec. sch.(mother)*$)$, $\mathbb{1}($*fem. HoH*$)$,*age HoH O(2),* $\mathbb{1}($*Some prim. sch. HoH*$)$, $\mathbb{1}($*Some sec. sch. HoH*$)$, $\mathbb{1}(>$ *sec. sch. HoH*$)$.									
Data: *DHS lifestyle*									
High B.P.	.0076[a]	0.001	.0108[a]	0.260	.244	6.17	475,157	15,838	1
	(.0008)		(.0008)		(.430)	(1.57)			
Asthma	0.00002	0.000	.00012	0.019	.015	6.18	712,978	15,546	1
	(.00012)		(.00012)		(.122)	(1.57)			
Diabetes	.0019[a]	0.001	.0015[a]	0.028	.014	6.19	677,232	15,545	1
	(.0001)		(.0001)		(.117)	(1.57)			
Obese	.0128[a]	0.006	.0100[a]	0.154	.077	6.07	851,767	28,330	38
	(.0005)		(.0003)		(.267)	(1.67)			
Controls: *age O(2),* $\mathbb{1}($*Some prim. sch.*$)$, $\mathbb{1}($*Some sec. sch.*$)$, $\mathbb{1}(>$ *sec. sch.*$)$, $\mathbb{1}($*fem. HoH*$)$,*age HoH O(2),* $\mathbb{1}($*Some prim. sch. HoH*$)$, $\mathbb{1}($*Some sec. sch. HoH*$)$, $\mathbb{1}(>$ *sec. sch. HoH*$)$.									
Data: *Afrobarometer*									
Fear Walking	.0157[a]	0.003	.0155[a]	0.430	.381	5.65	26,437	2,210	23
	(.0037)		(.0034)		(.486)	(1.76)			
Fear at Home	.0094[a]	0.001	.0102[a]	0.386	.334	5.65	26,437	2,210	23
	(.0037)		(.0036)		(.472)	(1.76)			
Theft at Home	.0042	0.000	.0059[b]	0.320	.288	5.65	26,476	2,210	23
	(.0028)		(.0026)		(.453)	(1.76)			
Attacked	.0026	0.000	.0024	0.147	.103	5.65	26,468	2,210	23
	(.0019)		(.0019)		(.303)	(1.76)			
Controls: $\mathbb{1}(<$ *Primary sch.*$)$, $\mathbb{1}($*Some sec. sch.*$)$, $\mathbb{1}(>$ *high sch.*$)$, *age O(2),* $\mathbb{1}($*fem.*$)$, *H.H. size*									

Note: *Regressions of respondent level 'outcome' on log population density in a 5km disk. Standard errors are clustered by 'survey cluster'. Each row reports results from two regressions, one without demographic controls and one with;* $^a = 1\%$, $^b = 5\%$, $^c = 10\%$, *all two-tailed tests. Relevant demographic controls are listed at the bottom of each panel.* $\overline{y}$ *and* $\overline{x}$ *are mean of outcome and* $\ln($*pop. density*$)$ *in the 'no-controls' sample. Except for the LSMS panel, we lose only a tiny number of observations when we add controls.*

Summary

With more complete data, moving to the city looks like a complicated trade-off.

- Better: Income, public utilities, status of women(mostly), innoculations.
- Worse: Domestic Abuse, Infant mortality, Childhood illness, lifestyle diseases, crime.

This suggests a story quite similar to the one we started with for the developing world. People move to cities for better wages, but face a worse disease environment. The difference is that new urbanites seem not to be working in manufacturing.
This suggests that ‘spatial equilibrium’ is relevant to this process, but it would be nice to understand migration costs better.

Glaeser, Edward. 2011. *Triumph of the City*. Pan.

Haines, Michael R. 2001. The urban mortality transition in the united states, 1800-1940. In *Annales de démographie historique*, 1. 33–64.

Henderson, J Vernon and Matthew A Turner. 2020. Urbanization in the developing world: Too fast, too slow or just right? .

of Agriculture, United States Dept. 2020. Quick stats. *National Statistics Service, Washington, D.C.* .