Big Data
for measuring the digital economy

Valeria Jordán

UNECLAC
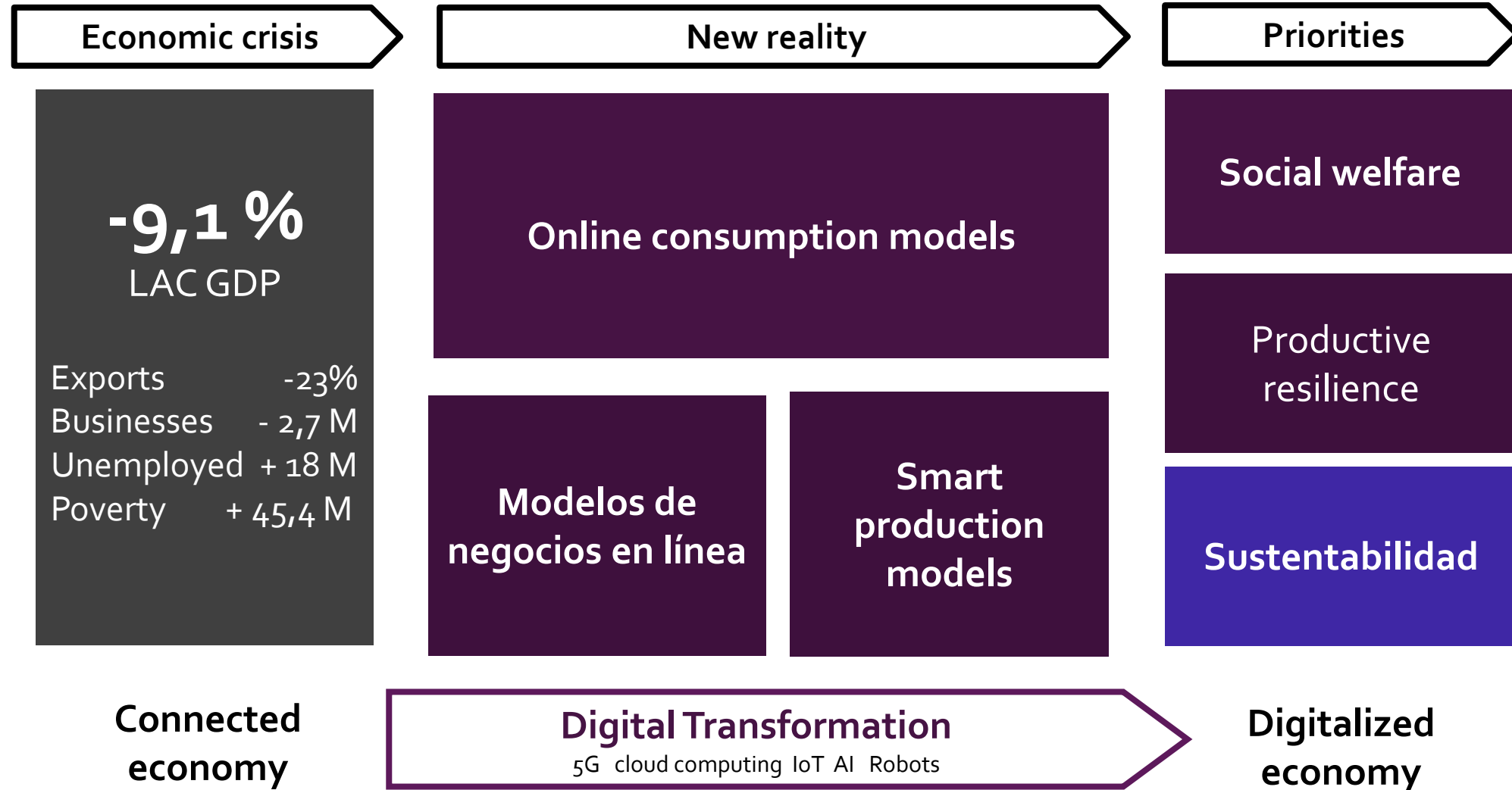
COVID-19 and the acceleration of digitization

Economic crisis

New reality

Priorities

-9,1 %
LAC GDP

Exports   -23%
Businesses   - 2,7 M
Unemployed   + 18 M
Poverty   + 45,4 M

Online consumption models

Modelos de negocios en línea

Smart production models

Social welfare

Productive resilience

Sustentabilidad

Connected economy

Digital Transformation
5G cloud computing IoT AI Robots

Digitalized economy

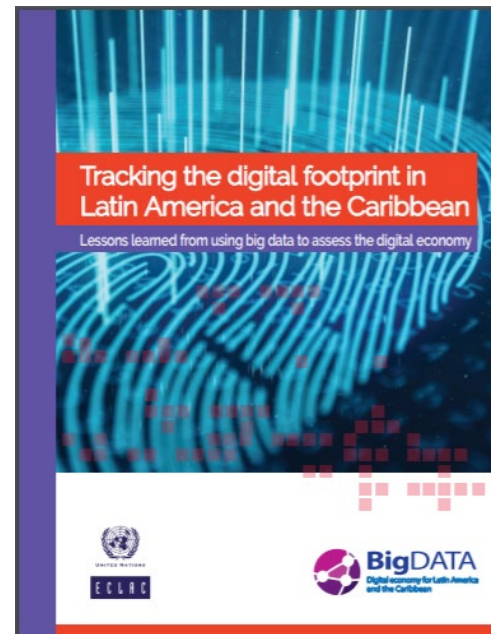# Why we need to innovate in measuring the digital economy?

- Insufficient data to understand a phenomenom with great economic and social impact

- Indicators are needed beyond connectivity (e-commerce, digitization of companies, gig-economy, fintech, etc.)

- The COVID-19 pandemic presented a greater need for data to identify trends (distance learning, remote working, e-health, etc)

- Digital footprints as new sources of data

  - Digital platforms: Social networks, marketplaces
  - Website's content
  - Electronic transactions
  - Call Detail Records
  - GPS

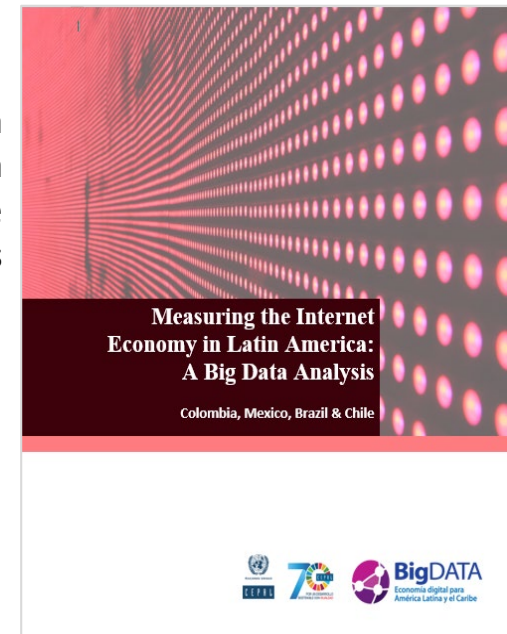# Big Data for measuring the digital economy in LAC countries

ECLAC project with funds from the UN Development Account (2016 – 2020)

- Objetive: improve national capacities in the Latin America and the Caribbean region to measure the digital economy using big data analytics and traditional statistics to support evidence-based policy design.

- Activities:

  i. Capacity building on big data techniques.

  ii. Experimental exercises to generate indicators

Web data

Web data combined with administrative records

Tracking the digital footprint in Latin America and the Caribbean

Lessons learned from using big data to assess the digital economy

Measuring the Internet Economy in Latin America: A Big Data Analysis

Colombia, Mexico, Brazil & Chile

Big Data techniques

Call Data Analysis

It is based on CDR which is an information structure that stores applicable information about a given telephone activity.

Benefit: access to a large amount of available data.

Disadvantage: requires agreements between parties

API (Application Programming Interface): access to databases provided by digital platforms

Benefit: quick and easy access to the data provided.

Disadvantage: at the mercy and trust of the API provider, may be biased

Web scraping/Web crawler: A technique to control the extraction of information from a site, in which specific data from the web is collected and copied, usually to a central local database or spreadsheet, for later retrieval or analysis.

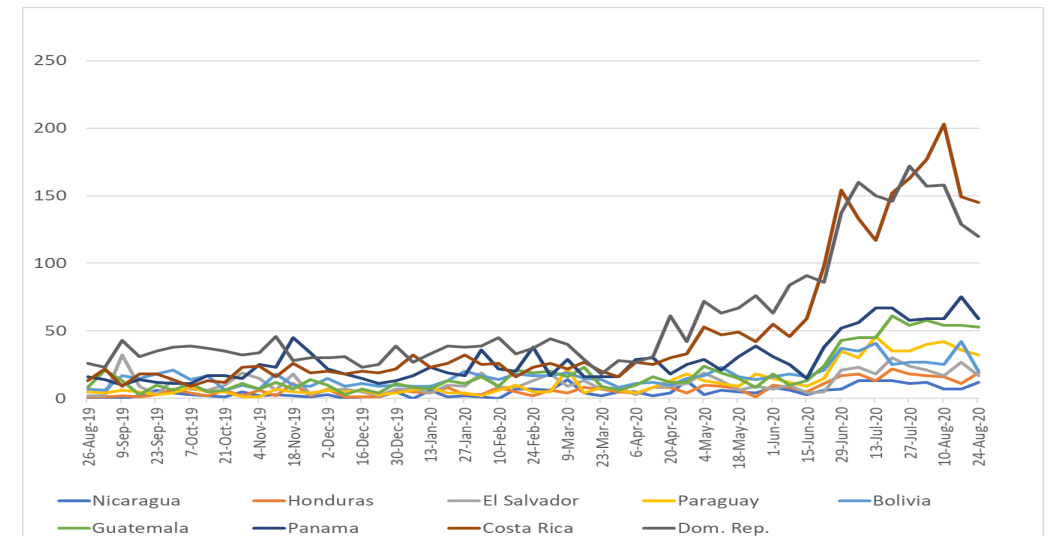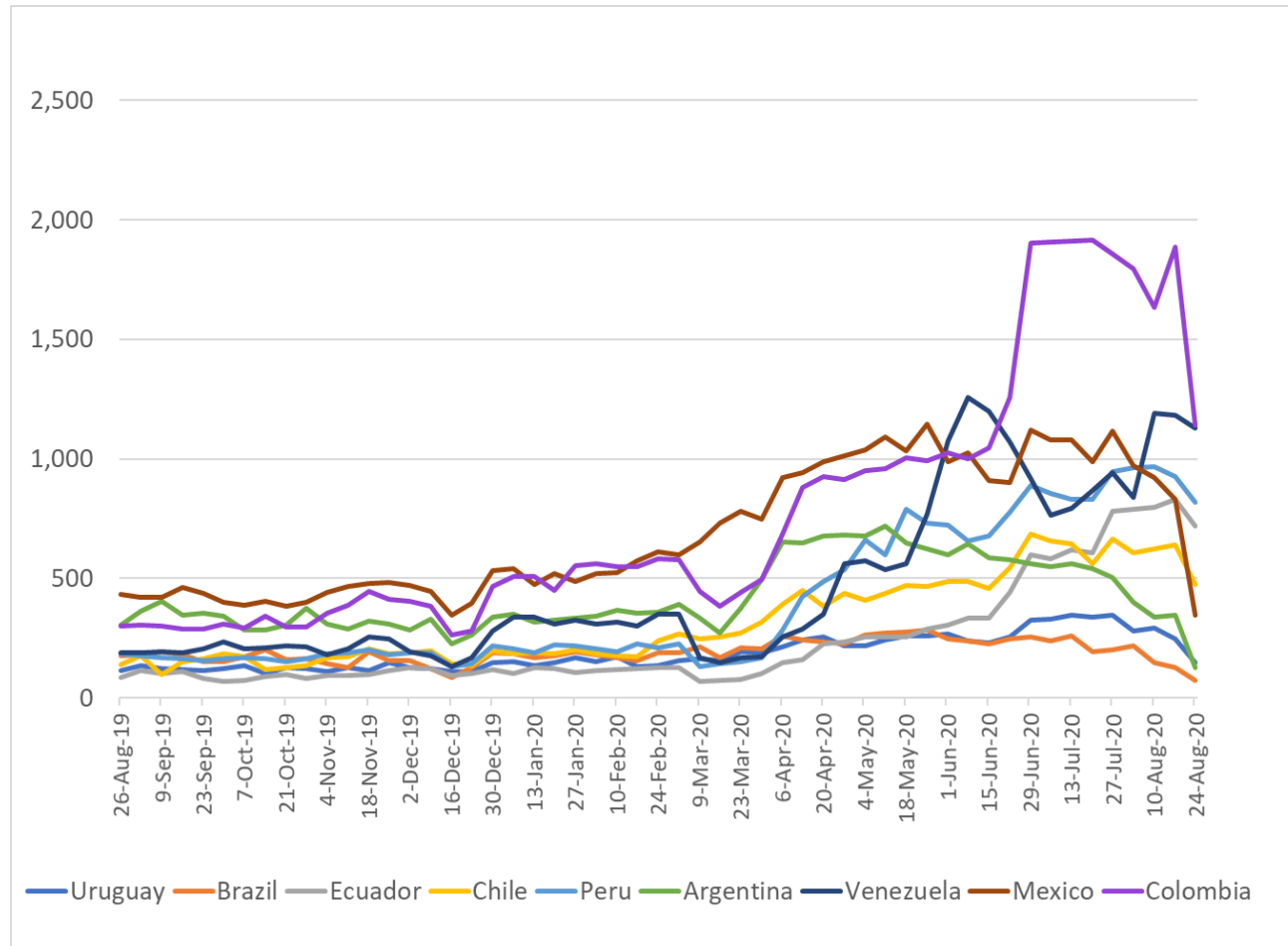Benefit: access to a large amount of data available to the public, updated in real time

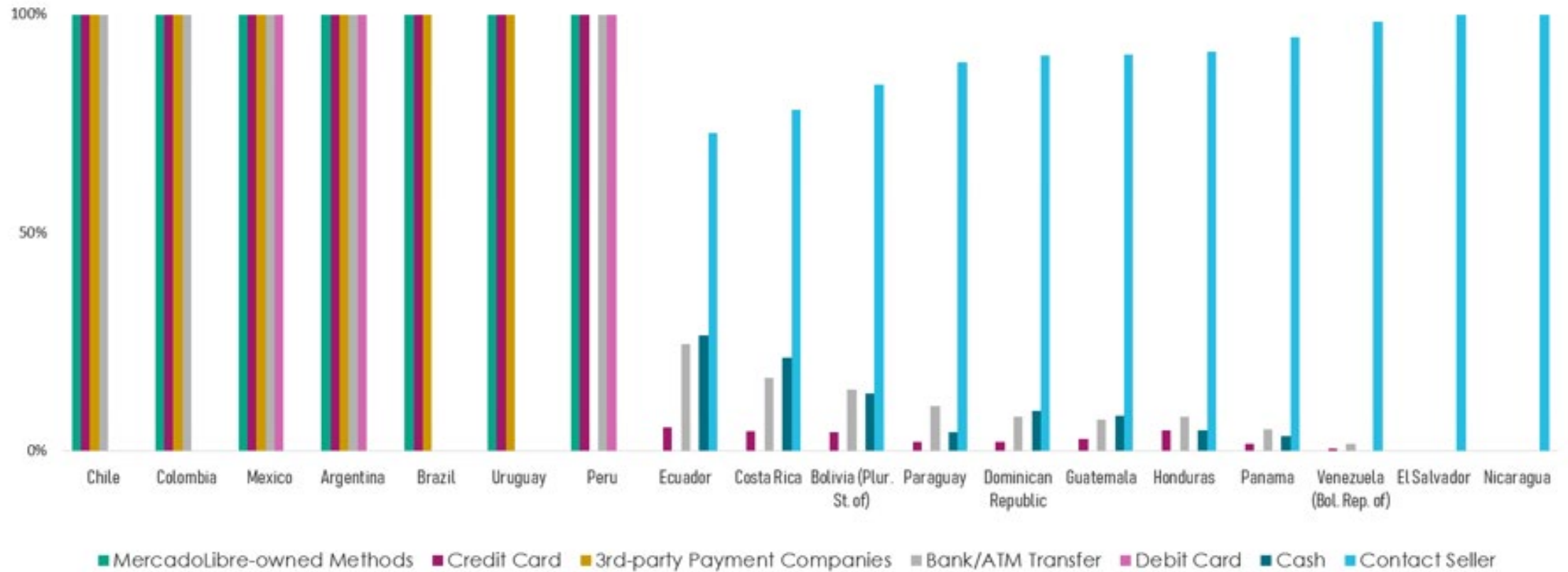Disadvantage: complex and resource-intensive work (computational).

Some results

Sellers on marketplaces

Number of MercadoLibre.com seller registrations by market size, 2019-2020
(units)

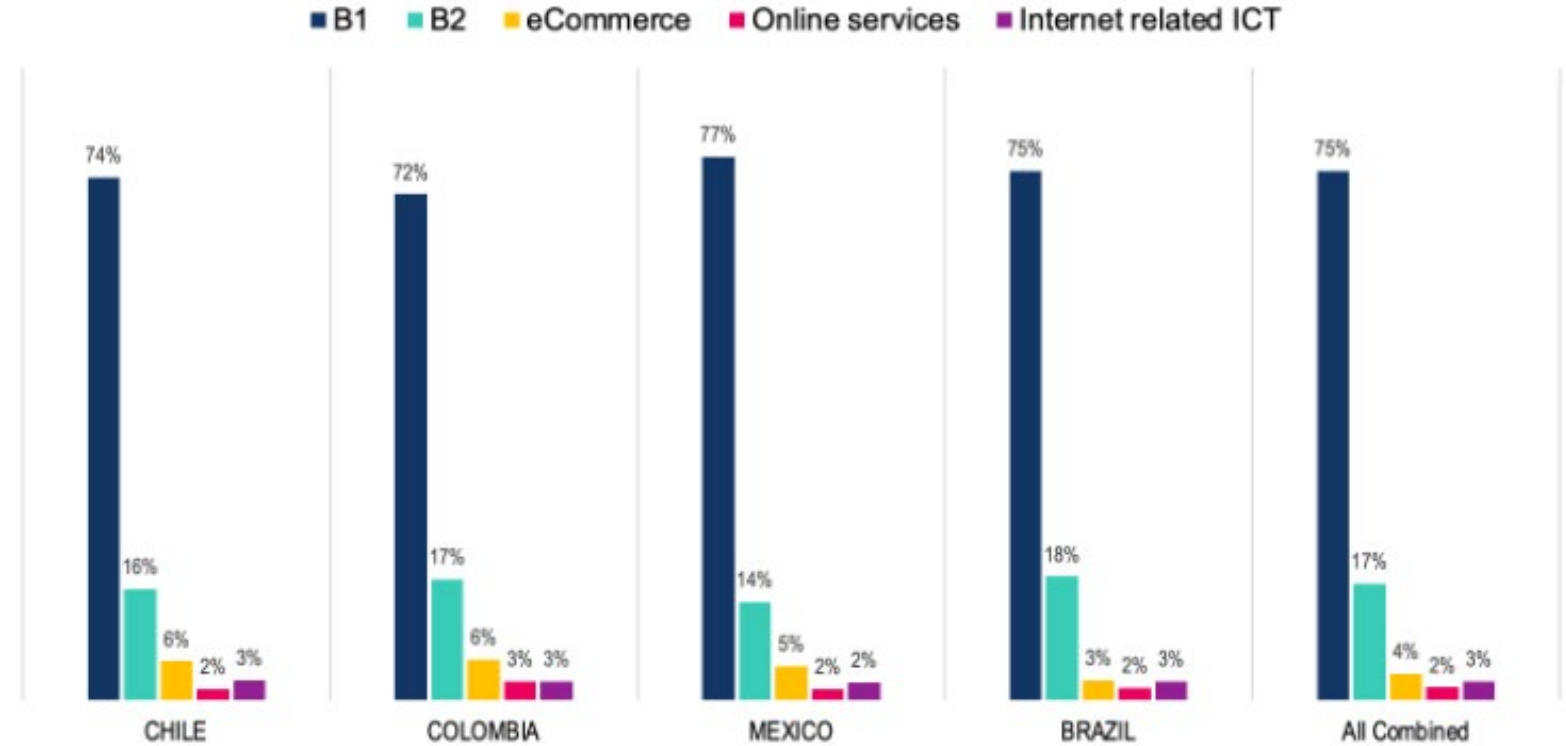MercadoLibre available payment options by country. January 2019

(In percentages)

Growth of online offer

50%

of all products offered on the e-commerce sites by PYMEs were published in a online catalogue between March – September 2020

Cumulative offered products in Sampled Shopify PYME product catalogues in selected countries per online publication date

Legend: Colombia · Chile · Brazil · Mexico

75% of businesses have passive presence

The core of the internet economy makes up 8% (Brazil) - 12% (Colombia) and about 50% belongs to online Stores

20% switched to transactional with the pandemic

Distribution of companies by type of online presence, March 2020
(percentage)

- B1
- B2
- eCommerce
- Online services
- Internet related ICT

CHILE: 74%, 16%, 6%, 2%, 3%

COLOMBIA: 72%, 17%, 6%, 3%, 3%

MEXICO: 77%, 14%, 5%, 2%, 2%

BRAZIL: 75%, 18%, 3%, 2%, 3%

All Combined: 75%, 17%, 4%, 2%, 3%

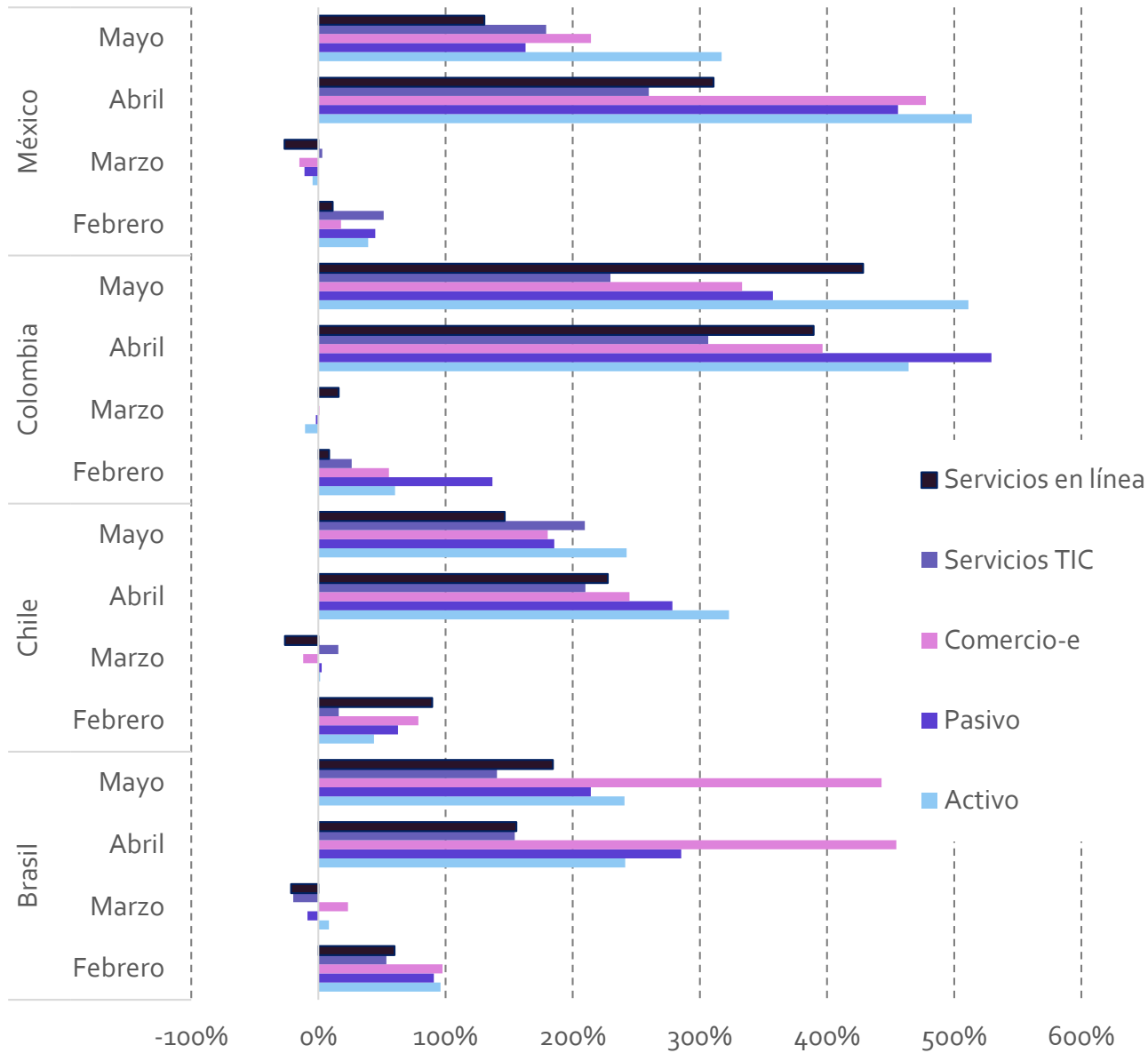Source: ECLAC, Project "Big Data for measuring the digital economy", 2020.

Businesses moved to online presence during the COVID-19 pandemic

Number of new business websites registered per month in selected countries and monthly growth
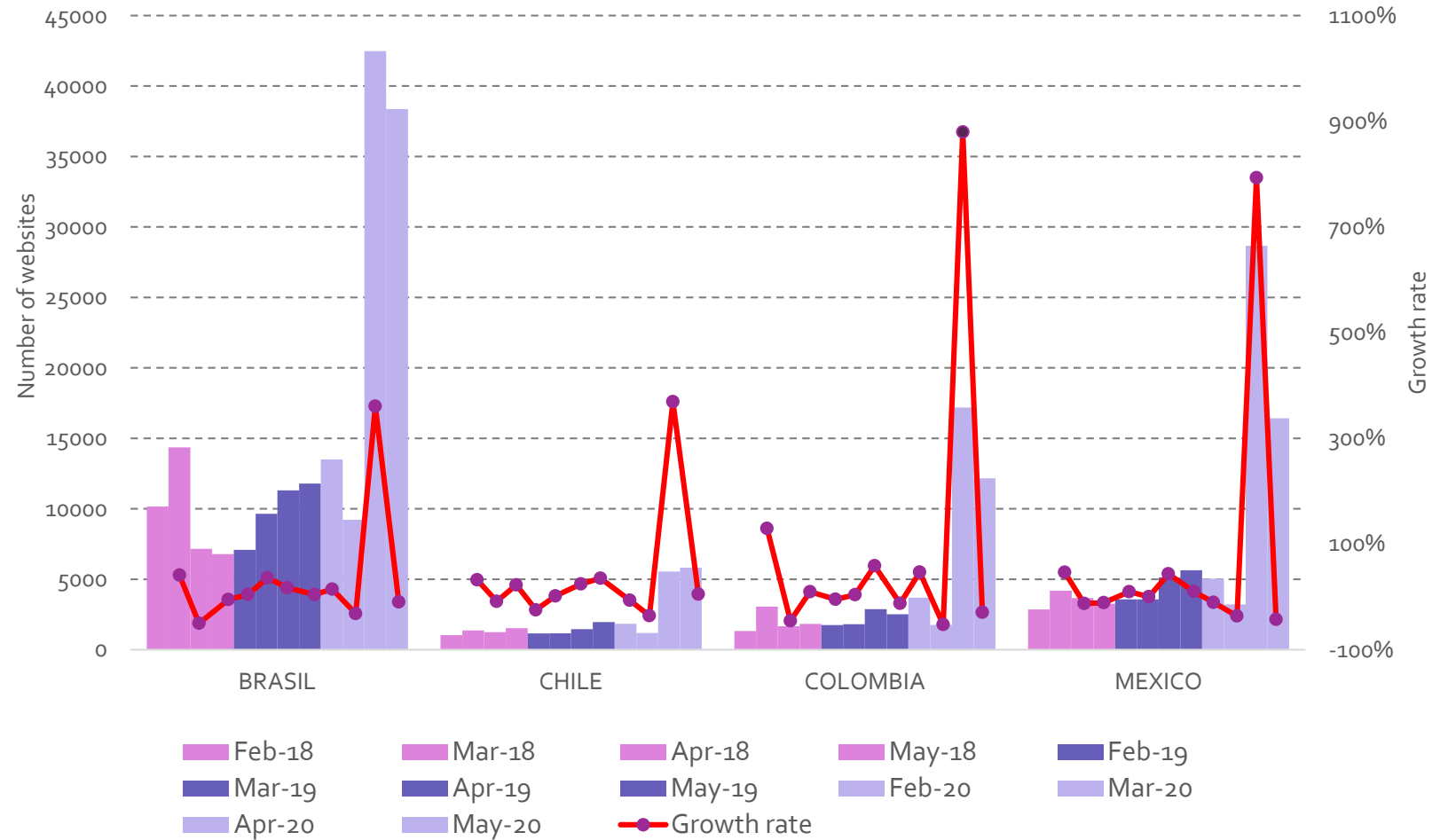
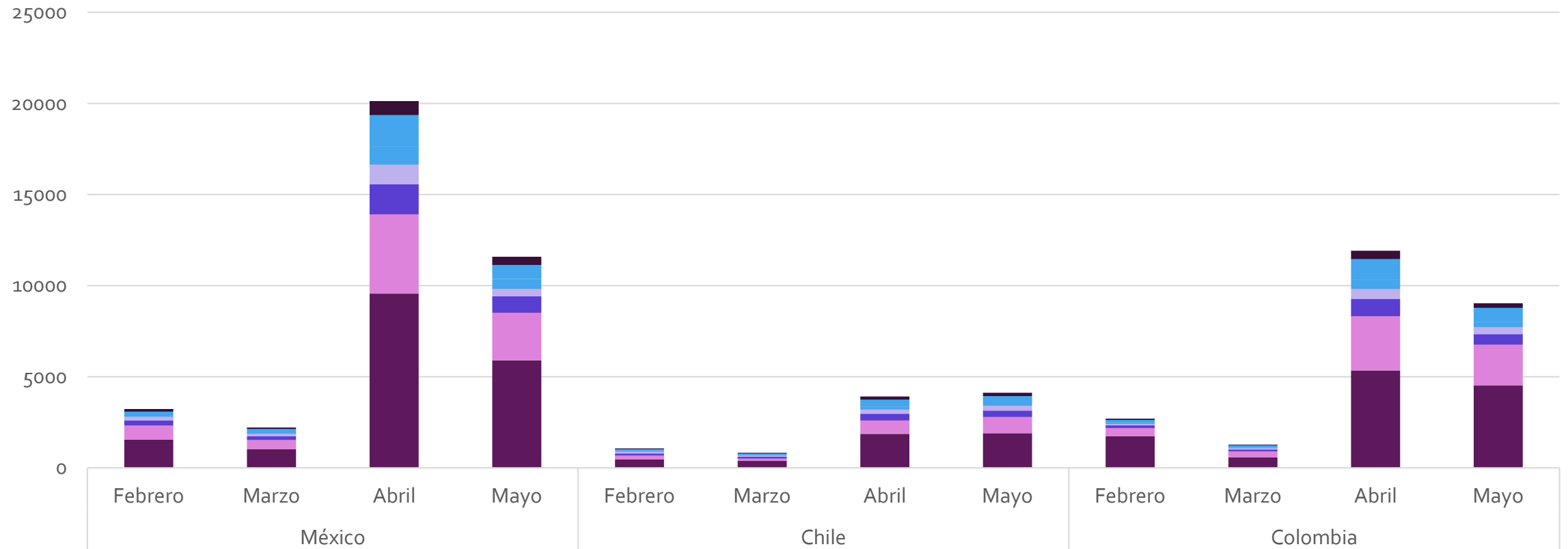Source: ECLAC, Project Big Data for measuring the digital economy".

Crecimiento interanual de sitios web empresariales por tipo en países seleccionados (en porcentajes)

Servicios en línea
Servicios TIC
Comercio-e
Pasivo
Activo

Explosion of e-commerce sites

Retail and businesses services go online

Source: ECLAC, Project "Big Data for measuring the digital economy", 2020.

More and more data
tech prices, crowdfunding, freelancers, labor market, rental and accommodation…
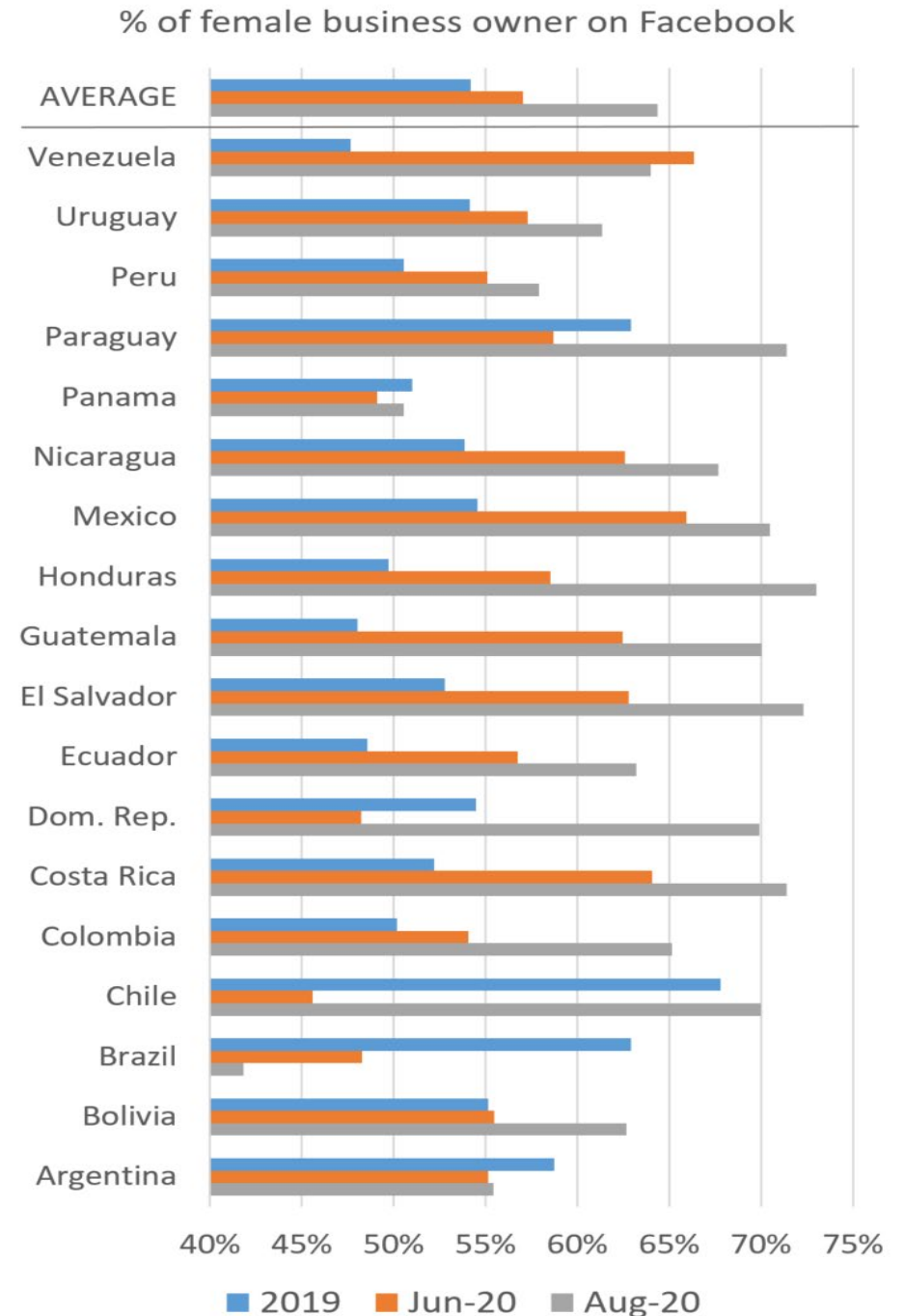
In FB during the pandemic, entrepreneurs have expanded :

Females X 3 (from 1.6 million to 4.1 millions)

Males X 2.5 (from 1.2 million to 3.0 millions)

% of female business owner on Facebook

The challenge of combining web data with business administrative records

Completeness of information for 6 variables used to merge the Business Registry to the website database. (in percentages) (Information based on Nov 2020)
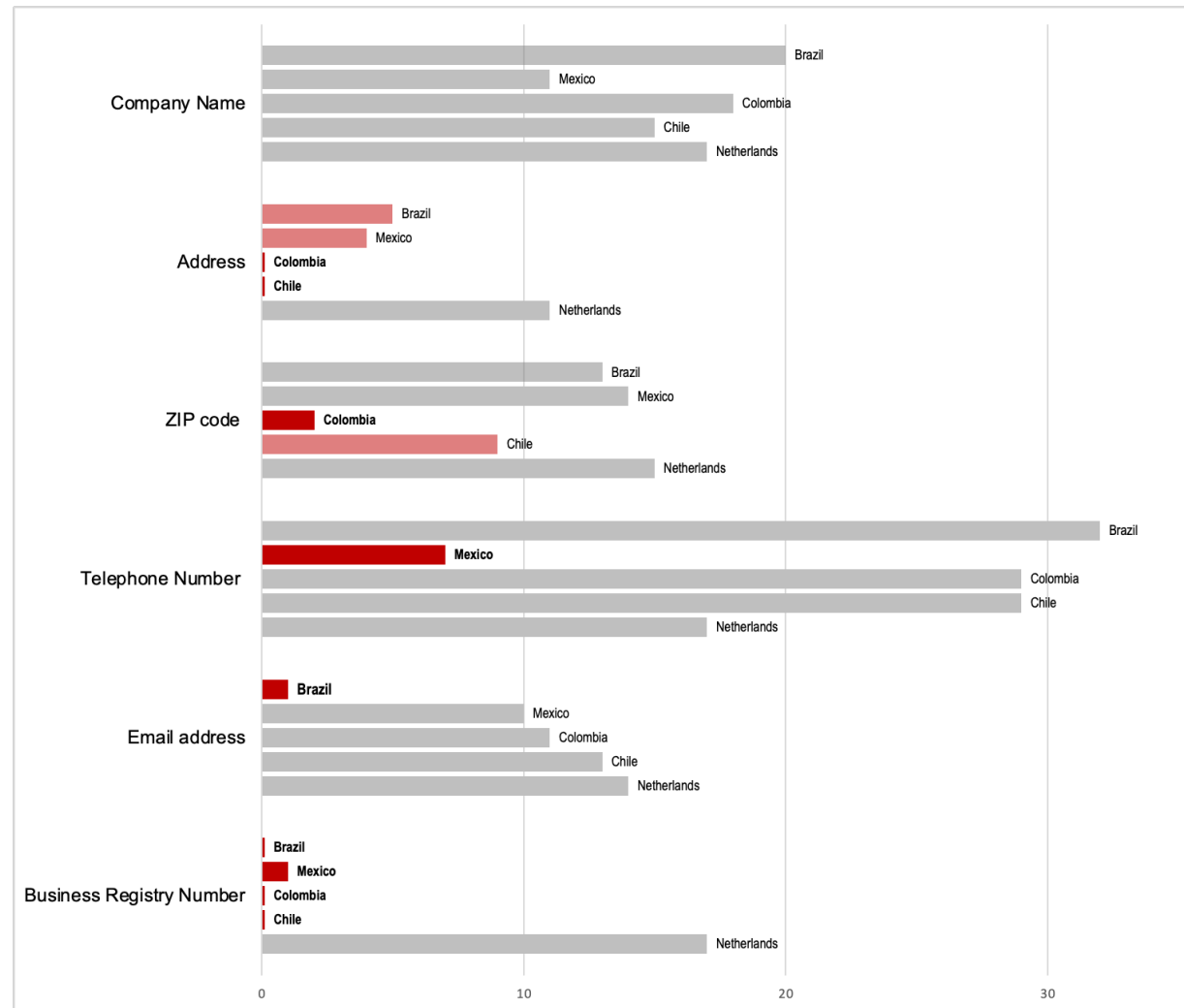
Web data can be used to generate indicators and improve administrative records

LESSONS

1. Unprecedented diversity of data that is useful to understand new paradigms

2. Accessing data is always difficult, there is no magic solution

3. Big data and traditional statistics are complementary: different purposes

Original   Administrative records   Big Data

4. Data Innovation requires institutional adjustments and new capabilities

THANKS!

FOR MORE INFORMATION

Big data for measuring the digital economy (valeria.jordan@cepal.org)

https://www.cepal.org/en/proyectos/big-data-grandes-datos-la-economia-digital-america-latina-caribe

COVID-19 Observatory in Latin America and the Caribbean

https://www.cepal.org/en/topics/covid-19