A MACHINE LEARNING APPROACH TO NOWCASTING THE GROWTH RATE OF THE ICT SECTOR

Molly Lesher
Head, Digital Policy, Economics, and Measurement Unit
OECD Directorate for Science, Technology and Innovation

5th International Seminar on Big Data for Official Statistics, 29-31 May 2024, Xiamen, China

OCDE
DES POLITIQUES MEILLEURES POUR UNE VIE MEILLEURE

Evidence-based digital policies require advancing the digital economy measurement agenda

- Digital transformation impacts every sector, but the ICT sector is at its core
- Lags in the production of sectoral growth rates make it hard to know if policies are achieving their desired goal
- Non-traditional data and machine learning techniques can serve as an important complement to official statistics
- Real-time estimates of how the ICT sector is performing today inform policy decisions that impact this vital sector in the future

An empirical and agnostic approach to measuring ICT sector growth

- The model only relies on two sources of data:
  - Annual estimates of ICT sector value added in volumes (OECD STAN Database)
  - Monthly Search Volume Index series of Google Trends' categories.
- Employs a data-driven approach to:
  - Examine current economic growth trends of the ICT sector in OECD countries
  - Compare them with overall economy-wide growth rates
  - Provide policy makers with up-to-date and cross-country comparable data on the economic growth of this key sector

An indicative matching of Google Trends categories related to the ICT sector

ISIC Rev.4 code	ISIC Rev. 4 division	Google Trends category	Google Trends category ID
D26	Computer, electronic and optical products	Computer Servers	728
		Consumer Electronics	78
		Binoculars, Telescopes, Optical Devices	1384
D61	Telecommunications	Telecom	13
		Mobile Phones	390
		Communications Equipment	385
D62-63/D62T63	IT and other information services	Network Storage	729
		Internet Software	807
		Web Services	302

Note: This table includes a non-exhaustive list of available Google Trends categories related to the ICT sector.

Sources: OECD STAN Database, *http://oe.cd/stan*; Google Trends (accessed on 19 February 2024).

Extracting useful information from Google Trends data

Name	Description	Correction(s)
Sampling noise	Google Trends data suffers from sampling noise, especially in smaller regions and less popular categories.	1st: use the average of five different samples. 2nd: exclude search indexes with a variance exceeding 10 across the five samples.
Downward trend	Google Trends data show a downward trend over time due to the increasing total search volume on the internet.	Apply Lowess smoothers (locally weighted scatterplot smoothing) to filter out the common trend to country specific SVI series.
Seasonality	Searches vary throughout the year and very often exhibit strong seasonal patterns.	1st: apply a Hodrick-Prescott filter, 2nd: incorporate seasonal dummy variables to account for any remaining seasonal patterns.

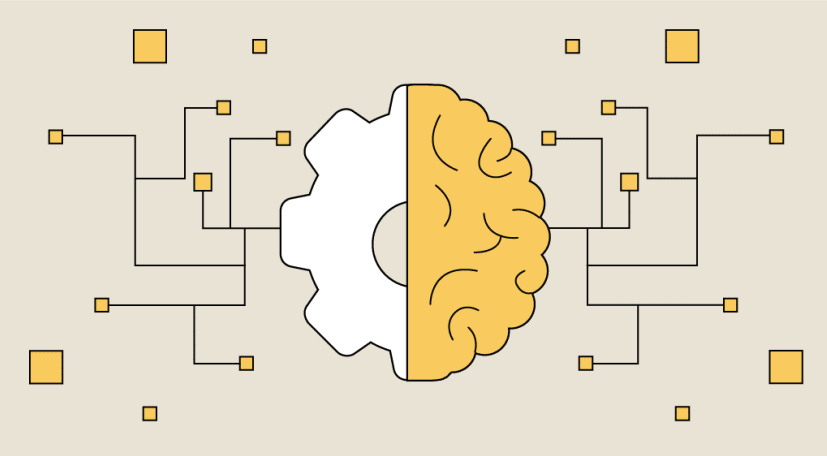
Choosing the best model – 1st step: statistical method selection

Table 5.1. Comparing different statistical methods

	Hodrick-Prescott Levels	Lowess Levels	Fixed Effects	Hodrick-Prescott Log	Lowess Log	Hodrick-Prescott + Lowess
RMSE training data	0.32	2.55	2.67	0.0001	0.05	2.53
RMSE validation data	2.71	2.92	2.86	3.41	2.89	2.58

Note: The table shows RMSE for the training and validation data for different statistical methods.
Source: Author's calculations using OECD STAN database and Google Trends data.

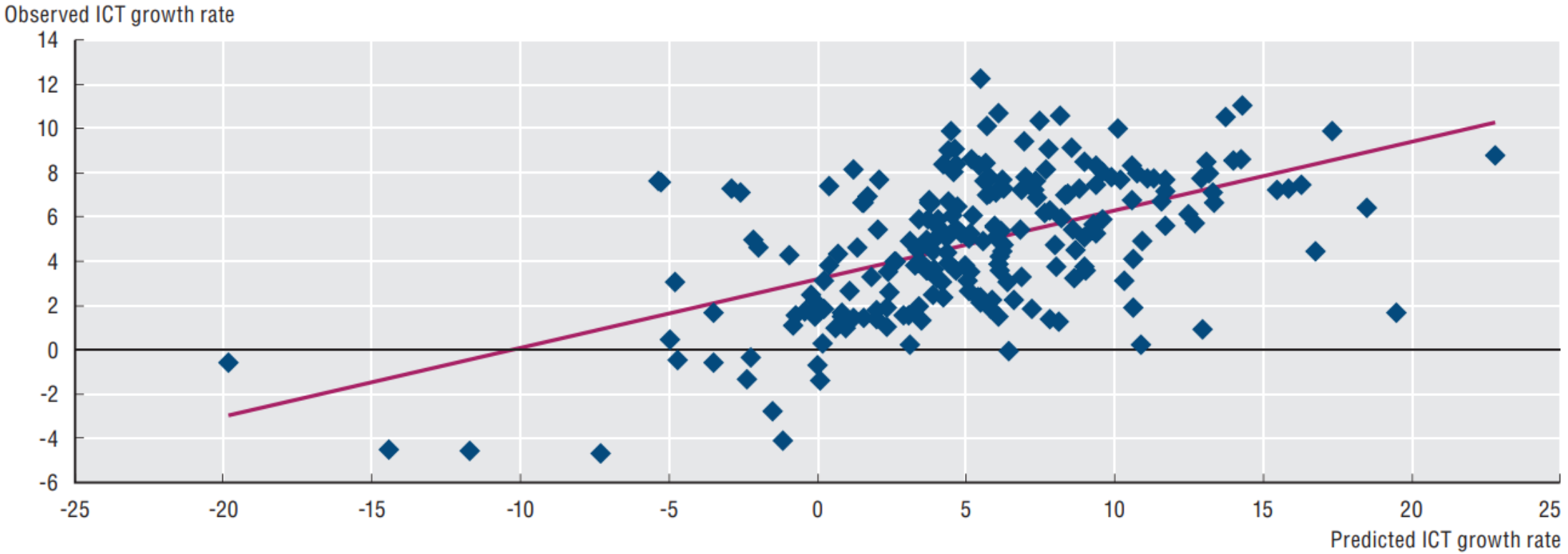
Choosing the best model – 2nd step: machine learning model selection



- Different machine learning methods were tested, including two stages, gradient boosting and neural networks
- AR(1) was also performed for benchmark comparison
- Best machine learning performance using artificial neural networks
- Choosing the best parameters using an automated process
- Calculating standard errors using a bootstrapping procedure similar as the one used for the OECD Weekly tracker of GDP

The nowcasting model performs strongly

Correlation between the observed and predicted ICT sector growth rates, 2011-19



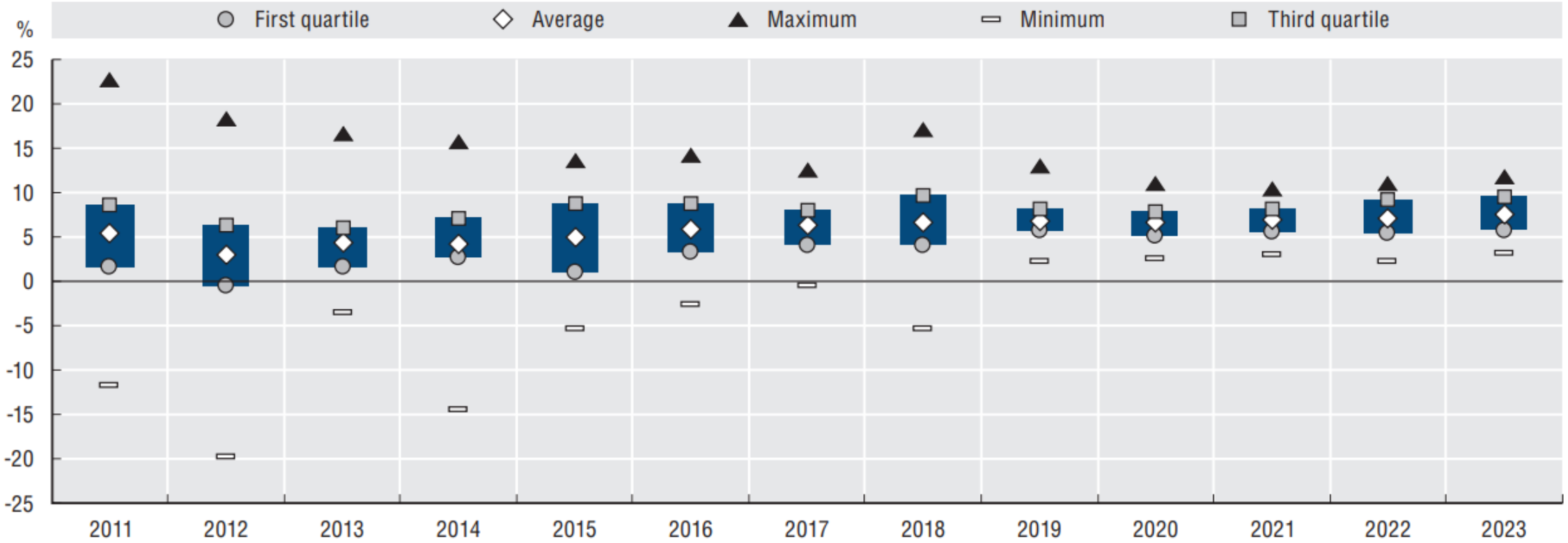
Notes: ICT = Information and communication technology. The correlation coefficient is 0.54, significant at a 1% level.

Source: Authors' calculations using OECD STAN Database and Google Trends data (accessed on 19 February 2024). StatLink contains more data.

StatLink **https://stat.link/y27lp3**

The ICT sector shows remarkable dynamism

ICT growth rate distribution (observed and predicted), 2011-23

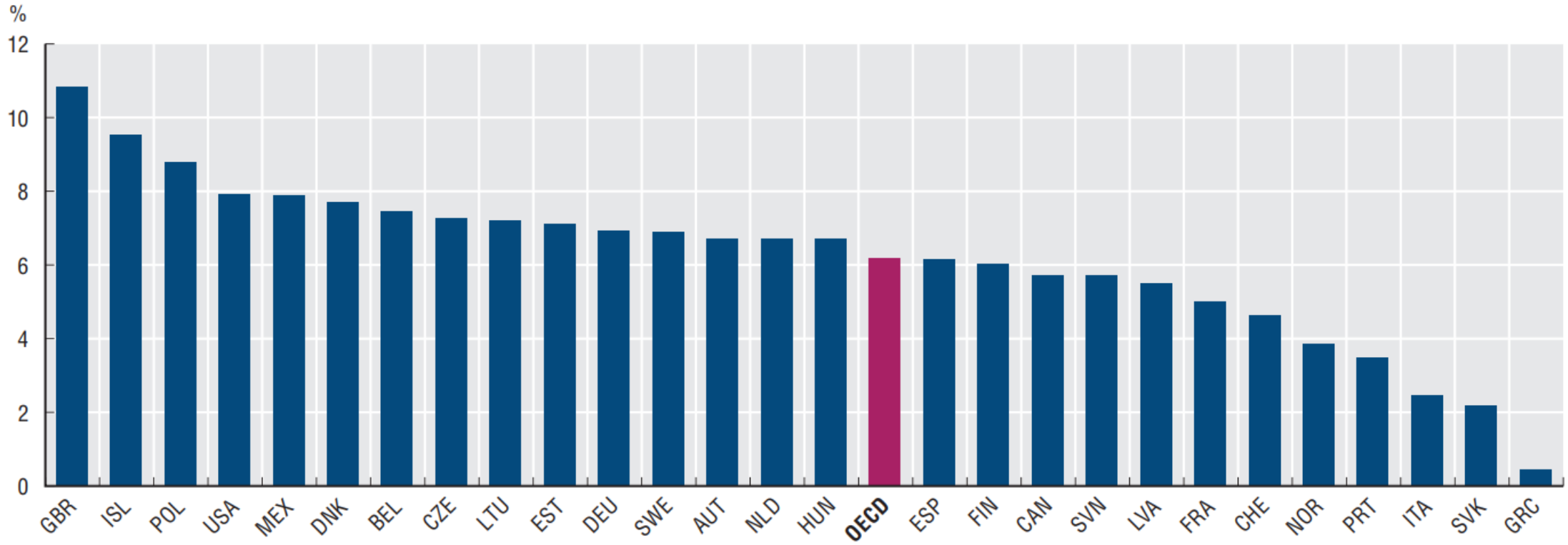


Source: Authors' calculations using OECD STAN Database and Google Trends data (accessed on 19 February 2024). StatLink contains more data.

StatLink **https://stat.link/rcbu6h**

In the past decade, the ICT sector grew in most OECD countries, but not equally

Average ICT sector growth rates (observed and predicted), 2013-23

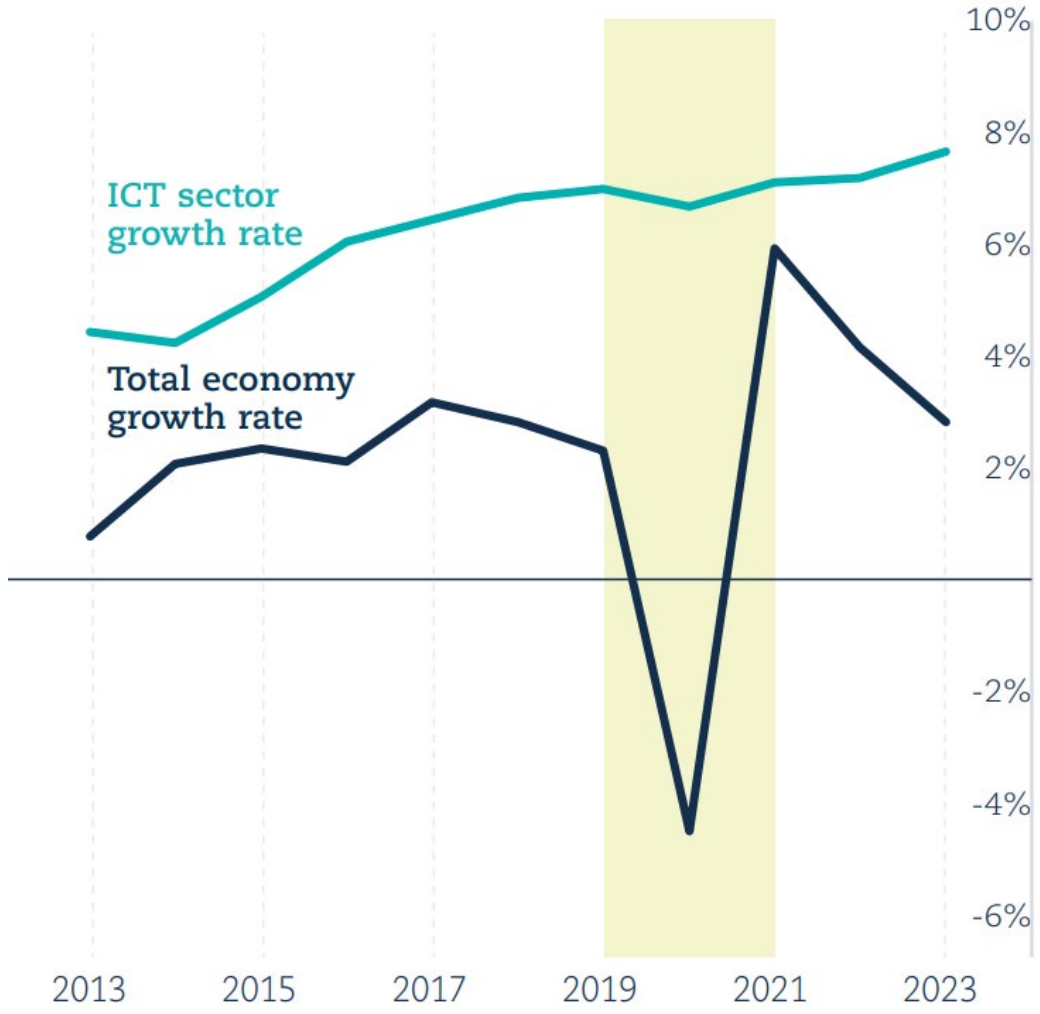


Note: This figure presents the mean observed and predicted ICT growth rates by country.

Source: Authors' calculations using OECD STAN Database and Google Trends data (accessed on 19 February 2024). StatLink contains more data.

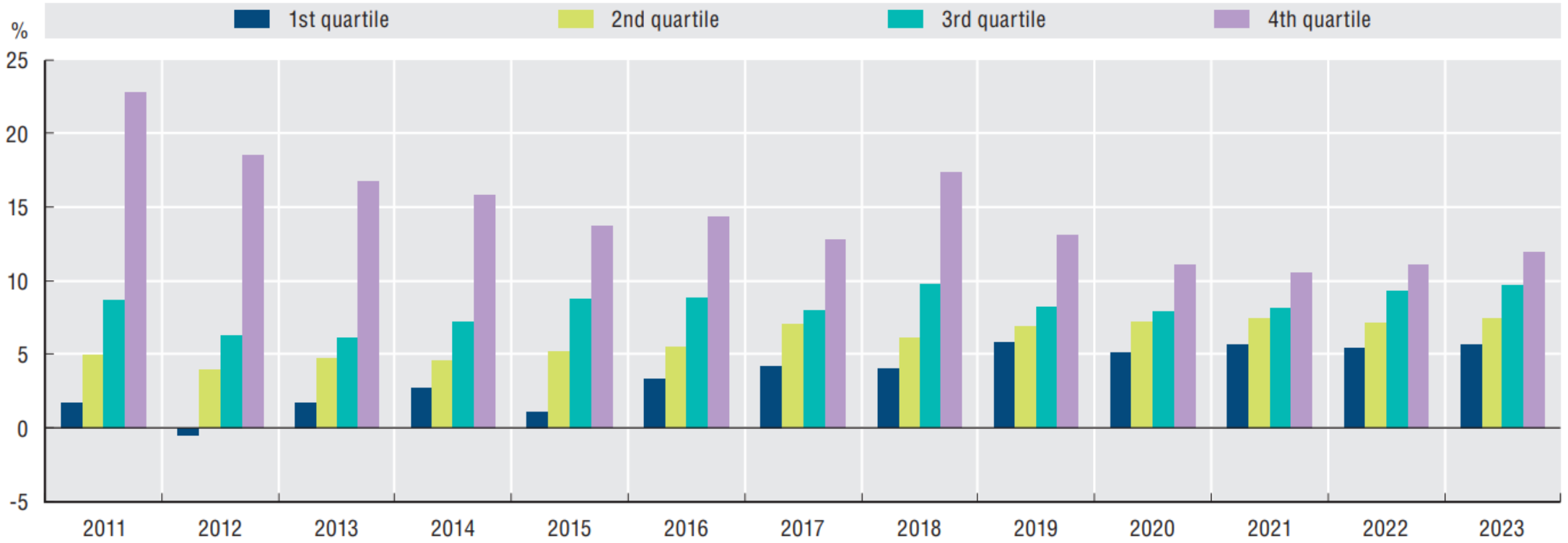
StatLink **https://stat.link/q3rcbp**

The ICT sector is resilient in the face of economic headwinds



ICT sector growth rates are converging across countries

Average ICT sector growth rates (observed and predicted) by quartile, 2011-23

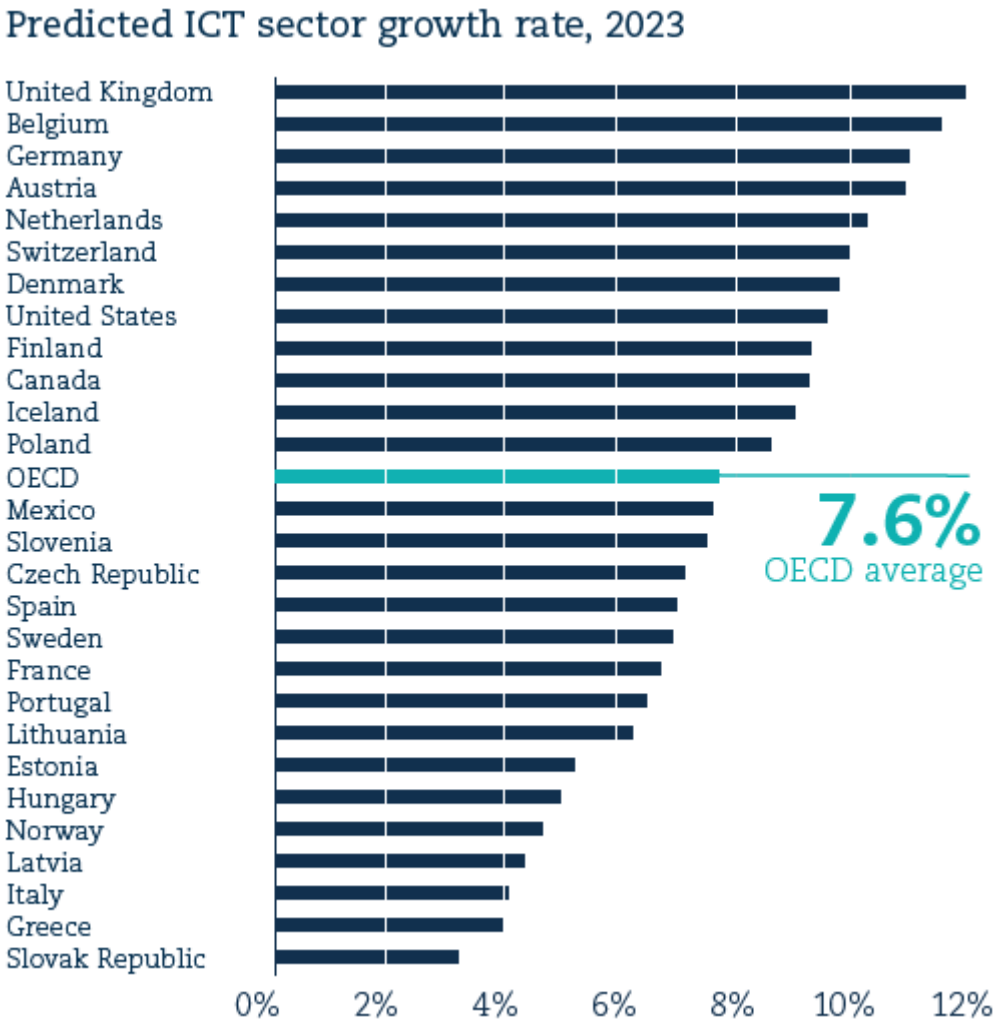


Notes: This figure presents the distribution of observed and predicted ICT growth rates from 2011 to 2023 by quartile. Values from 2011 to 2018 (earliest) originate from the STAN Database depending on the country, while those from 2019 (or 2020) to 2023 are nowcast estimates.

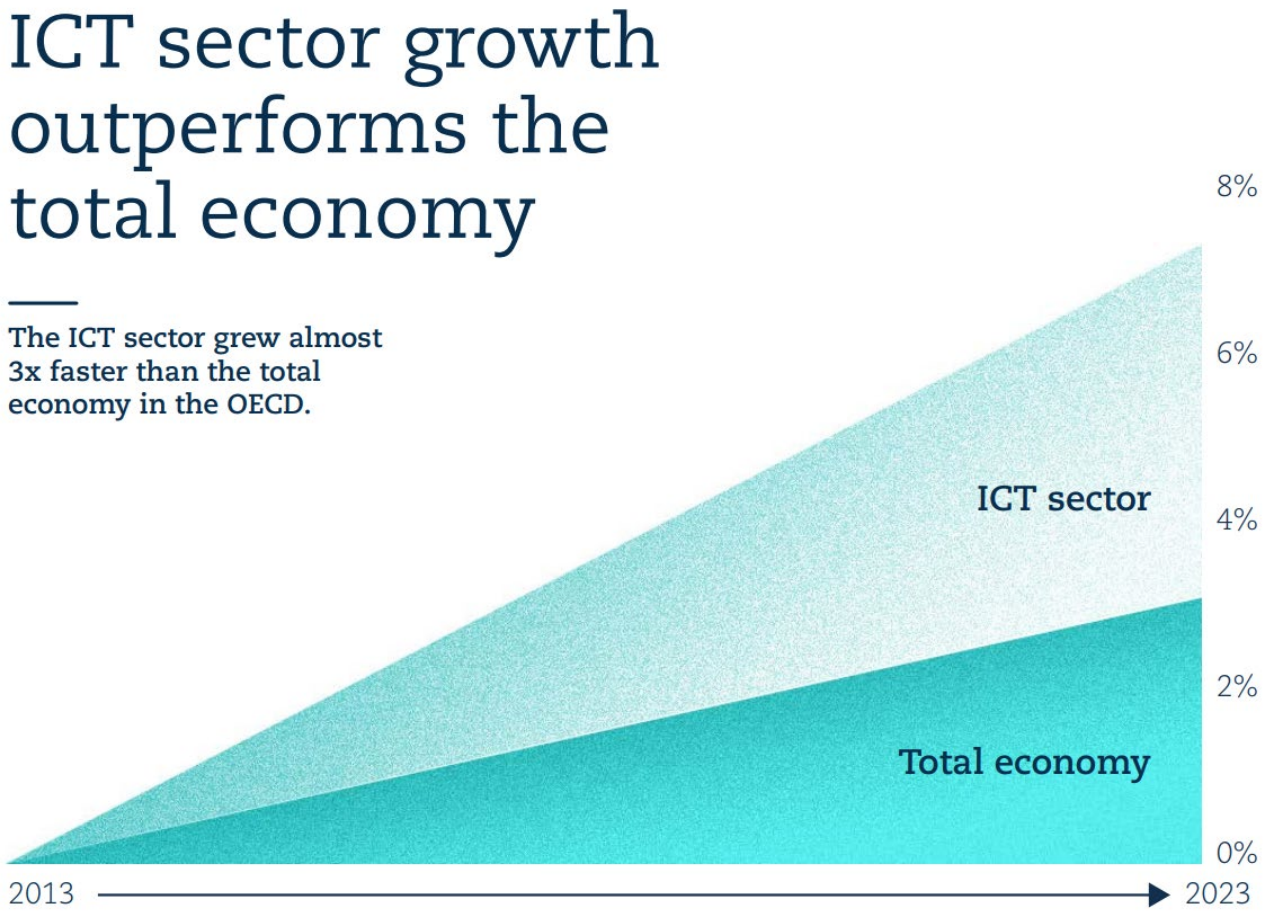
Source: Authors' calculations using OECD STAN Database and Google Trends data (accessed on 19 February 2024). StatLink contains more data.

StatLink **https://stat.link/tdpoly**

ICT sector growth was strong across countries in 2023

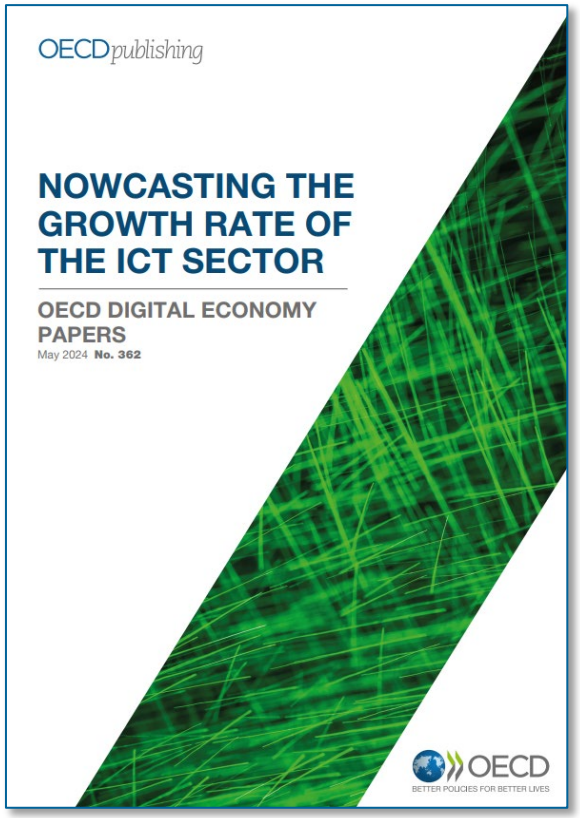


The ICT sector grew almost 3 times faster than the total economy over the past decade



Want to learn more?

https://www.oecd.org/publication/digital-economy-outlook/2024/



https://doi.org/10.1787/eb4938a0-en

Thank you for your attention

For any questions, please contact:

molly.lesher@oecd.org

OCDE
DES POLITIQUES MEILLEURES POUR UNE VIE MEILLEURE