Data Enabled Change Accelerator

Using UK tax data to predict businesses with high growth potential

Dr Rannia Leontaridi

Director of Business Growth, BEIS

What is the problem we are trying to solve?

Business data in government comes from a range of sources but is only accessible to particular staff for specific purposes.

Addressing the barriers to accessing and using business data would drive a wide range of benefits for Government and beyond.

Better policy and better implementation – making the UK the best understood economy.

IMPROVED ACCESS TO BUSINESS DATA

Targeting Support

Sharing data between departments to **identify businesses to support**.

Collaborating in supporting businesses using data insights to **improve take-up**.

Monitoring Progress & Evaluating Policy

Determining the **degree of success of Government programmes and policies**

Research & Analysis

Generating insights into patterns of business growth, productivity, sectoral and local trends, segmentation, to **drive policymaking**.

How are we tackling this problem?

The UK tax office (HMRC) has the richest source of admin data on business performance.

However HMRC data is protected by primary legislation. In our project we are testing the limits of what can be achieved within the law, working in partnership with HMRC.

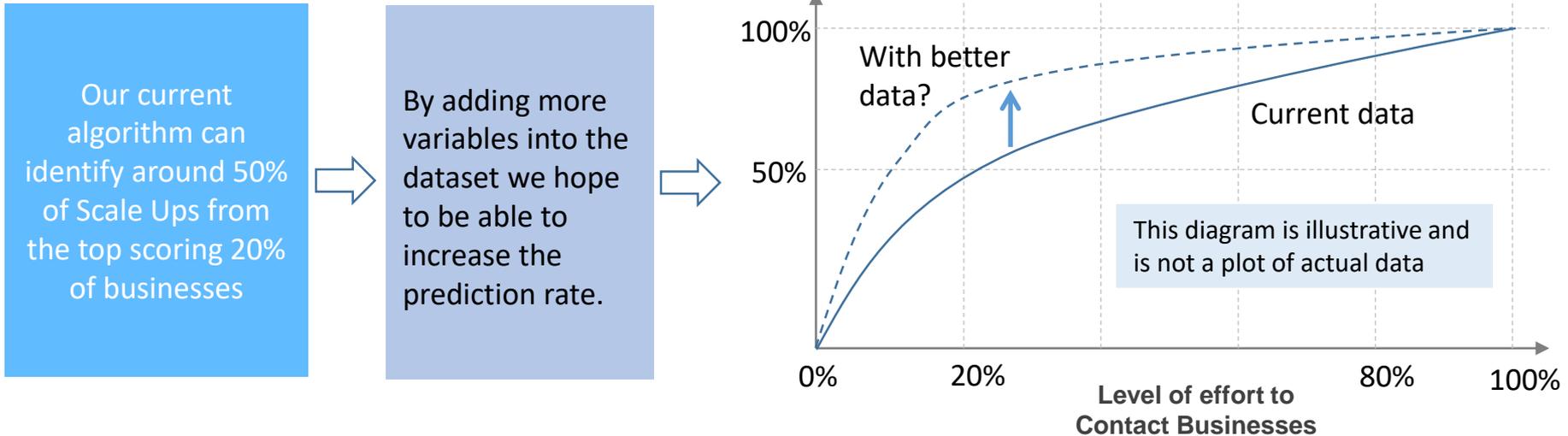We tax data that we have created includes records for **3.2 million UK businesses** and **over 200 variables**. Variables include: employment, growth, productivity, R&D credits, exports, profits, Directors.

We have so far identified **13,700 'Scale Up' businesses** in the dataset, with turnover growth rates of 20%+ per annum over a 3-year period.

BEIS has been working with the Office for National Statistics to develop this project, with input from a steering group of private sector and academia partners.

We are using data science techniques to predict high growth

We have developed an algorithm that can identify businesses with high growth potential, with a high degree of success.

Our current algorithm can identify around 50% of Scale Ups from the top scoring 20% of businesses

By adding more variables into the dataset we hope to be able to increase the prediction rate.

Identification Rate of Scale Up Businesses (jobs measure)

100%

With better data?

Current data

50%

This diagram is illustrative and is not a plot of actual data

0%   20%   80%   100%

Level of effort to Contact Businesses

Previous academic work, using less granular data, has suggested that high growth episodes are essentially **random** at the firm level, even if you could pick out some broader factors associated with growth.

Next we are launching operational trials, to test how to drive up engagement with business support to boost growth and productivity

We are about to launch messaging pilots, contacting hundreds of businesses in three areas of the UK:

- Leeds City Region

- West of England

- Derby, Derbyshire, Nottingham and Nottinghamshire

We will contact the businesses that the algorithm identified as having "high growth potential" to put them in touch with local support services for Scale Up businesses.

We have developed the communication with input from the Behavioural Insights Team, and we will be evaluating the outcome of the trials before launching further trials.

Later in 2019 we will trial targeting support at businesses with **low productivity**.

Our longer term vision is to embed this approach across government to support a step change in how we interact with business