Thursday 11 November

11.15-12.45: Session 1: Defining and Measuring Intangibles

Chair: Sylaja Srinivasan (Bank of England)

Josh Martin (Office for National Statistics) “The F words: why surveying businesses about intangibles is so hard”


Thomas Niebel (ZEW Mannheim) “Intangible Capital Indicators Based on Web Scraping of Social Media”
The F words: why surveying businesses about intangibles is so hard

Josh Martin
Head of Productivity, UK Office for National Statistics (ONS)
Twitter: @JoshMartin_ONS

11 Nov 2021 – at IARIW-ESCoE intangibles conference
Investment in Intangible Assets survey (UK 2009)

• Series of questions on each asset (6)
• Repeated in 2010
• Replicated in Italy and elsewhere in subsequent years
Survey estimates contradict macro estimates

Investment in intangible assets, UK market sector, £billions, 2008

Source: Martin and Baybutt (2021), based on ONS (2021) and Awano et al. (2010)
Why is collecting data on intangibles so hard?

• Intangibles are intangible!

• Data might not be held, or might be very dispersed

• Concepts unfamiliar to respondents

• Often created ‘on own-account’

• Valuation is difficult
Why is collecting data on intangibles so hard?

- Intangibles are easily *Forgotten*
- Definitions are *Fuzzy*
- Answers depend on question *Framing*
- Importance of survey *Frequency*
Forgotten
Businesses report inconsistently on intangible investment across surveys

Proportion of respondents inconsistent between surveys (IIA and UKIS), contingent on positive response in either survey, UK

Source: Martin and Baybutt (2021), based IIA1, IIA2, UKIS09 and UKIS11 microdata
Fuzzy
Survey estimates contradict macro estimates

Investment in intangible assets, UK market sector, £billions, 2008

Source: Martin and Baybutt (2021), based on ONS (2021) and Awano et al. (2010)
Different survey questions and guidance

Annual Business Survey (ABS)

Computer software programs and databases

i) developed by own staff for business use
ii) purchased or developed externally (bespoke)

Include:
- Program descriptions, extensions, supporting materials for systems and applications

Exclude:
- Hardware. Report this at 9 (e)
- Cost of ongoing management
Different survey questions and guidance

Quarterly acquisitions and disposals of Capital Assets Survey (QCAS)

Computer software programs

.include: program descriptions, extensions and supporting materials for systems and applications.

i) developed by own staff for business use?

ii) purchased or developed externally (bespoke)?
Different survey questions and guidance

Investment in Intangible Assets survey (IIA)

Software
During the reporting period, what was your business's expenditure on software bought from other organisations?

Include:
• off-the-shelf software
• databases
• software licences and licence renewals
• generic and bespoke software.

Exclude:
• software embedded in other items of current or capital expenditure, e.g. software preinstalled on IT hardware

During the reporting period, what was your business's expenditure on software development carried out by its own staff?

Include:
• staff costs of all staff involved, excluding contractors
• associated costs, including office facilities, overheads and materials but not capital items.

Note: Estimates based on proportions of staff time are acceptable
Different survey questions and guidance

UK Innovation Survey (UKIS)

During the 3 year period 1 January 2014 to 31 December 2016, did this business invest in any of the following, for the purposes of current or future innovation?

**Acquisition of advanced machinery, equipment and software for innovation**

- Advanced machinery and equipment
- Computer hardware
- Computer software

For each of the main innovation related investments in question 4, please **ESTIMATE** the amount of expenditure for the YEAR 2016 ONLY. Include both internal costs and purchases from outside the business.
Framing
The correlation of reported investment is higher within BERD than compared to UKIS

Correlation coefficients on investment in R&D, between UKIS and BERD, and between adjacent years of BERD, UK

Source: Martin and Baybutt (2021), based on various UKIS and BERD microdata

Office for National Statistics
Frequency
Persistent differences between respondents in quarterly and annual surveys

Percentage difference in average reported investment between annual (ABS) and quarterly (QCAS) surveys, selected assets, UK

- Purchased software
- Machinery and equipment
- Buildings and structures
- Own-account software

Source: Martin and Baybutt (2021), based on various QCAS and ABS microdata

Office for National Statistics
Designing a better intangibles survey
Designing a better intangibles survey

• Organisation by ‘data type’ (not by asset)
• Guidance on likely location of data
• Collect more ‘components’ rather than ‘totals’
• Provide more clarity (where helpful)
• Allow flexibility (in timings and methods)
• Solidify definitions (in research community)
• Do more testing!
Section F - Design of new or improved products or services

20. During the year, did this business fund any external or internal design of new or improved products or services?
Activities solely relating to the design of new or improved products or services. For example, design of blueprints, plans, patterns, original drawings, models and prototypes. Exclude work conducted as part of R&D and included in Section E.

Yes ☒ → Go to question 21
No ☒ → Go to question 24 20

21. During the year, what was this business’s expenditure on activities by other organisations to design new or improved products or services?
Include costs of bought-in design services. Exclude costs of design embedded in other items of current or capital expenditure.

£[__[__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__][__]
Thank you for listening!

Josh Martin
Head of Productivity, ONS
Topic Lead, ESCoE

Email: Josh.Martin@ons.gov.uk
Twitter: @JoshMartin_ONS
Discussant: Gaganan Awano (Cabinet Office)
Deflating Intangible Investment: Some New Ideas and Estimates

IARIW-ESCoE Special Conference on Measuring Intangibles
RSA House, London, UK
November 11, 2021

Leonard Nakamura, Federal Reserve Bank of Philadelphia, Emeritus

*The views expressed today are my own and not necessarily those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.*
We are in an extraordinary time for innovation and therefore for R&D, the quintessential intangible

- Innovative firms are rapidly expanding our capabilities
  - Apple, Google, Tesla, SpaceX, Netflix, Spotify, Illumina, Moderna, and Amazon
  - Creativity is at the heart of the modern corporation

- Extraordinary innovation in all dimensions of human endeavor
  - Software, statistics, data, art, travel, entertainment, biology, global culture, chemistry, education, sensing, imaging, electricity, communication, neurology, health, and markets.

- When we study intangibles, we venture into the very heart of creativity.
Intangibles differ from tangibles on many dimensions

• We are trying to tame intangibles. We try to view intangible investments as very similar to tangible investments.
  – But can we? Should we?
• Dimensions along which intangibles differ from tangibles
  – Nonrival and nonlinear—zero cost of reproduction
  – IP protection (sometimes refused); open source
  – No physical deterioration, no geographical ties
  – Not arm’s length and very risky with fat tails
  – Never the same from period to period
Pricing intangibles from their outcomes?

• In this paper, I focus on the deflation of research and development
  – By and large, we deflate R&D by deflating input costs for capital and labor and add back in an overall estimate of productivity.

• Generally, we do not observe transaction prices for R&D
  – We observe transaction prices for the products that are created
  – The key innovative products we study are both outputs of and inputs into innovation

• Tangible investments like a computer or communication equipment can fall rapidly in price
  – why not intangible investment?
Evidence of rapid improvement in R&D from depreciation rates

- R&D does not deteriorate physically over time
- Instead, it loses private value because of obsolescence or because it loses intellectual property protection
- Either way, obsolescence arises from technological progress, which remains permanent
- R&D depreciates fairly rapidly, often 16% annually or more
- This is one set of evidence that progress is fast in intangible investment and perhaps intangibles prices are deflating
Depreciation rates of R&D are often faster than 16 percent (modal rate of depreciation for R&D)

Table 1. Domestic Research and Development, 2018, Total and Selected Groupings, National Center for Science and Engineering Statistics (NCSES)

<table>
<thead>
<tr>
<th>Business Activity</th>
<th>Industry codes</th>
<th>Billions of dollars</th>
<th>Percentage</th>
<th>Depreciation rate spread in percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td>441.0</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>Medical and chemical</td>
<td>32500,33910,62150</td>
<td>98.7</td>
<td>22.4</td>
<td>9 to 16</td>
</tr>
<tr>
<td>Machinery and electronic</td>
<td>33300</td>
<td>103.4</td>
<td>23.4</td>
<td>25 to 40</td>
</tr>
<tr>
<td>Transportation machinery</td>
<td>33600</td>
<td>49.1</td>
<td>11.1</td>
<td>7 to 31</td>
</tr>
<tr>
<td>Information</td>
<td>51000</td>
<td>94.7</td>
<td>21.5</td>
<td>16 to 33</td>
</tr>
<tr>
<td>Professional, scientific and tech</td>
<td>54100</td>
<td>47.2</td>
<td>10.7</td>
<td>16</td>
</tr>
<tr>
<td>Subtotal of selected groupings</td>
<td></td>
<td>393.1</td>
<td>89.1</td>
<td>7 to 40</td>
</tr>
</tbody>
</table>
Rapid price declines for experimentation

- The cost of performing an experiment has fallen dramatically in many key instances:
- Space launches, robot scientists, cloud computing, neurology, biochemistry, sensors, AI, and on and on.
Examining central drivers of innovation we see very fast rates of price decline

• Central dogma of biology:
  – DNA instructs RNA how to create proteins, proteins fold into 3 dimensional shapes to function

• DNA sequencing of one human genome:
  – 2005: $10 million each; 2018: $1000 each; 47 % annual rate of price decline

• DNA manipulation:
  – 2012 to 2018: fall in price 150 fold; even faster than DNA sequencing

• New Deep Mind AI technology solves protein folding for 2/3 of all proteins!
DNA sequence fall in costs
Other examples

- Moore’s Law: 29% rate of price decline
- Cloud computing prices: 7% price decline
- Rocket cost per flight fell by 3 times: 13%
- LEDs: 21%
- Batteries: 20%
- Lidar (laser equivalent of radar): 22%
- Internet per byte: 28%
- Cellular per byte: 45%
Rapid rates of price decline characterize central innovative products

• Price declines to outputs of R&D include input prices as well as output prices:
  – Computation, data storage, DNA sequencing, DNA manipulation, protein folding, communications, rockets, robots, AI, sensors, imaging

• Both depreciation rates and direct measures of price declines suggest that many areas of R&D are experiencing deflation

• If the outcomes of R&D fall in price rapidly, is that deflation of R&D

• Do we need better measures of price changes in intangible investment as we do in consumer products?
Conclusion

- There has been rapid progress in some areas of intangible investing
- It is possible that some areas of research and development could be deflating at double digit rates
- Whether we should use rapid deflation for some intangible investment is an open question
  - If we want to draw a strong parallel with tangible investing, perhaps we should deflate intangibles
Discussion of *Deflating Intangible Investment: Some New Ideas and Estimates*

Marshall Reinsdorf, discussant

IARIW-ESCoE Conference on Measuring Intangible Assets and their Contributions to Growth
RSA House, London
November 11-12, 2021
Capturing the Gains from Technological Progress by Tracking Declines in the Cost of Achieving Outcomes

Nordhaus (1996) tracked the cost of achieving an outcome -- lumens -- to measure the effect of advances in lighting technology.

Costs of achieving outcomes like sequencing the human genome, manipulating DNA, launching a spacecraft, performing a chemistry experiment, training an AI algorithm, ..., have fallen dramatically.

Large price declines in equipment/software embodying intangible assets imply rapid obsolescence of intangible assets, and therefore high depreciation rates.


Links between Depreciation, Price Change, and Growth of Investment

Paper doesn’t calculate $/outcome analogs to $/lumen, but the cost must have fallen fast

Rapidly falling deflator would imply rapid growth of investment in intangible assets

High depreciation rate then needed to avoid exaggeration of the growth of the stock

Rapidly growing gross investment is often associated with high rates of retirement of old assets

When prices of a capital asset are falling fast, we should suspect that its depreciation rate is high
Comments

Nice paper, with an important insight and lots of fascinating examples of technology advances that may be fast enough to qualify as one of Nordhaus’ tectonic shifts!

Tectonic shifts in technology are hard to predict and irregular, which seems to fit the description of “Other changes in volume of assets” in the SNA

*Ex post* loss of value not suitable for measuring depreciation: the rate of obsolescence should reflect *expected* loss of value with aging and can’t be volatile

But faster rates of depreciation could be assumed to account for the frequency of big advances
For tangible assets, the age-price profile gives the rate of depreciation

- Used asset’s price relative to the price of a similar new asset gives the rate of depreciation
- The price index tracks the change over time in the price of a new asset

Paper should discuss how to recover these two pieces of information in the intangibles case

Yesterday’s reservation price of the asset embodying the new technology gives the price index, while today’s price of the asset embodying the old technology gives the rate of obsolescence
INTANGIBLE CAPITAL INDICATORS BASED ON WEB SCRAPING OF SOCIAL MEDIA

Patrick Breithaupt (ZEW), Reinhold Kesler (UZH), Thomas Niebel (ZEW) and Christian Rammer (ZEW)

IARIW-ESCoE Conference on Intangible Capitals
11 November 2021
RSA House London
MOTIVATION: KNOWLEDGE-BASED CAPITAL

- Knowledge-based capital is considered as a key factor for productivity growth (and dispersion).

- Measurement of knowledge-based capital is challenging.
- While the first innovative property, computerised information, are covered by different statistical surveys, comprehensive statistical data on economic competencies are scarce.
MOTIVATION: KNOWLEDGE-BASED CAPITAL

Can be measured by three main components according to Corrado et al. (2005, 2009):

1. Computerized information
2. Innovative property
3. Economic competencies

We focus on brand equity & firm-specific human capital

<table>
<thead>
<tr>
<th>Asset type</th>
<th>Included in National Accounts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computerized information</td>
<td></td>
</tr>
<tr>
<td>1. Software</td>
<td>Yes</td>
</tr>
<tr>
<td>2. Databases</td>
<td>Yes</td>
</tr>
<tr>
<td>Innovative property</td>
<td></td>
</tr>
<tr>
<td>3. Mineral exploration</td>
<td>Yes</td>
</tr>
<tr>
<td>4. R&amp;D (scientific)</td>
<td>Yes</td>
</tr>
<tr>
<td>5. Entertainment and artistic originals</td>
<td>Yes</td>
</tr>
<tr>
<td>6. New product/systems in financial services</td>
<td>No</td>
</tr>
<tr>
<td>7. Design and other new product/systems</td>
<td>No</td>
</tr>
<tr>
<td>Economic competencies</td>
<td></td>
</tr>
<tr>
<td>8. Brand equity</td>
<td></td>
</tr>
<tr>
<td>a. Advertising</td>
<td>No</td>
</tr>
<tr>
<td>b. Market research</td>
<td>No</td>
</tr>
<tr>
<td>9. Firm-specific resources</td>
<td></td>
</tr>
<tr>
<td>a. Employer-provided training</td>
<td>No</td>
</tr>
<tr>
<td>b. Organizational structure</td>
<td>No</td>
</tr>
</tbody>
</table>

Source: Corrado et al. (2017)
OBJECTIVES/CONTRIBUTION

- Develop a new way of measuring investments in economic competencies that do not require firm surveys.
- **Advantage**: broader coverage at substantially lower costs, a much higher timeliness, and a much higher frequency.

I. Develop a method for (semi-automated) linking of business surveys with platform data.

II. Develop non-monetary indicators that use publicly available data from online platforms and are related to *economic competencies*:
   - Facebook: **brand equity (# of likes)**
   - Employer branding and review platform Kununu: **brand equity & firm-specific human capital (ratings for company image & on-the-job training/career development)**

III. Validate the two indicators through a micro-data comparison with the MIP data (German contribution to the Community Innovation Surveys - CIS).
METHODOLOGY: IDENTIFYING PLATFORM PROFILES

1. Get Information on URL from company survey
2. Standardized Google search: "Info of the company + Name of the platform + site:URL of the platform"
3. Get Information on URL from company survey
4. Download page of first search result
5. Link profiles based on URLs or company names
6. Merge platform data with company survey
7. Check matches
   - automatically
   - manually
8. Obtain valid platform profiles

Mannheim Innovation Panel (MIP 2017)
EXAMPLE SEARCH AND MATCH: KUNUNU PLATFORM

<table>
<thead>
<tr>
<th>Name</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAP SE</td>
<td><a href="http://www.sap.de">www.sap.de</a></td>
</tr>
</tbody>
</table>

Google search for "www.sap.de Kununu site:www.kununu.com"

KUNUNU score: 4.3, Recommendation: 94%

- Working atmosphere: ★★★★★
- Work-life balance: ★★★★★
- Career / further education: ★★★★★
IDENTIFIED PLATFORM PROFILES AND MATCHING CRITERIA

- Google search for Facebook and Kununu for each of the 7,498 companies results in

Facebook
  - Exact match of the company URL (on platform)

Kununu
  - exact string matching (no exact match established: perform a fuzzy string matching)

2,114 (28.2%)

1,516

598

941

1,539 (20.5%)
### DESCRIPTIVE STATISTICS - ESTIMATION SAMPLES

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training: Kununu rating</strong></td>
<td>519</td>
<td>3.31</td>
<td>3.38</td>
<td>0.79</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td><strong>Training expenditures (MEUR)</strong></td>
<td>519</td>
<td>1.10</td>
<td>0.060</td>
<td>13.8</td>
<td>0.00097</td>
<td>300</td>
</tr>
<tr>
<td><strong>Turnover (MEUR)</strong></td>
<td>519</td>
<td>404.2</td>
<td>27.5</td>
<td>2671.4</td>
<td>0.080</td>
<td>46800</td>
</tr>
<tr>
<td><strong>Number of employees</strong></td>
<td>519</td>
<td>1238.3</td>
<td>165</td>
<td>8614.6</td>
<td>1</td>
<td>122608</td>
</tr>
<tr>
<td><strong>Number of Kununu ratings (Training)</strong></td>
<td>519</td>
<td>22.9</td>
<td>9</td>
<td>74.3</td>
<td>4</td>
<td>1174</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Image: Kununu rating</strong></td>
<td>492</td>
<td>3.62</td>
<td>3.75</td>
<td>0.80</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td><strong>Marketing expenditures (MEUR)</strong></td>
<td>492</td>
<td>6.98</td>
<td>0.11</td>
<td>81.4</td>
<td>0.0010</td>
<td>1480</td>
</tr>
<tr>
<td><strong>Turnover (MEUR)</strong></td>
<td>492</td>
<td>312.7</td>
<td>27.4</td>
<td>1765.2</td>
<td>0.22</td>
<td>25763</td>
</tr>
<tr>
<td><strong>Number of employees</strong></td>
<td>492</td>
<td>1014.4</td>
<td>158</td>
<td>7189.5</td>
<td>3</td>
<td>122608</td>
</tr>
<tr>
<td><strong>Number of Kununu ratings (Image)</strong></td>
<td>492</td>
<td>21.7</td>
<td>8.50</td>
<td>72.7</td>
<td>4</td>
<td>1171</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Image: Facebook likes</strong></td>
<td>944</td>
<td>9683.8</td>
<td>224</td>
<td>98292.4</td>
<td>1</td>
<td>1702502</td>
</tr>
<tr>
<td><strong>Marketing expenditures (MEUR)</strong></td>
<td>944</td>
<td>2.40</td>
<td>0.032</td>
<td>48.8</td>
<td>0.00048</td>
<td>1480</td>
</tr>
<tr>
<td><strong>Turnover (MEUR)</strong></td>
<td>944</td>
<td>58.2</td>
<td>5</td>
<td>434.9</td>
<td>0.027</td>
<td>11630</td>
</tr>
<tr>
<td><strong>Number of employees</strong></td>
<td>944</td>
<td>218.4</td>
<td>40.5</td>
<td>1090.8</td>
<td>1</td>
<td>25247</td>
</tr>
</tbody>
</table>
EMPIRICAL APPROACH

Standard OLS regression to analyze the relationship between the new indicator **brand equity** (company image) and the MIP 2017 survey-based measures on **marketing**:

\[
\ln Y_{image/likes,i} = \beta_{exp} \ln (expenditures\_marketing)_{2016,i} + X_i \gamma + e_i
\]

Standard OLS regression to analyze the relationship between the other new indicator **firm-specific human capital** and the MIP 2017 survey-based measures on **training expenditure**:

\[
\ln Y_{training,i} = \beta_{exp} \ln (expenditures\_training)_{2016,i} + X_i \gamma + e_i
\]
## OLS REGRESSIONS KUNUNU

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>(1) Training: Kununu rating</th>
<th>(2) ln(Training: Kununu rating)</th>
<th>(3) Image: Kununu rating</th>
<th>(4) ln(Image: Kununu rating)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Training expenditures)</td>
<td>0.0680*</td>
<td>0.0237*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.81)</td>
<td>(1.88)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Marketing expenditures)</td>
<td></td>
<td>0.0842***</td>
<td>0.0272***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.13)</td>
<td>(3.19)</td>
<td></td>
</tr>
<tr>
<td>ln(Turnover)</td>
<td>0.0121</td>
<td>0.0103</td>
<td>-0.0407</td>
<td>-0.00946</td>
</tr>
<tr>
<td></td>
<td>(0.28)</td>
<td>(0.68)</td>
<td>(-0.83)</td>
<td>(-0.61)</td>
</tr>
<tr>
<td>ln(Number of employees)</td>
<td>-0.0590</td>
<td>-0.0238</td>
<td>-0.0459</td>
<td>-0.0160</td>
</tr>
<tr>
<td></td>
<td>(-1.06)</td>
<td>(-1.27)</td>
<td>(-0.78)</td>
<td>(-0.83)</td>
</tr>
<tr>
<td>Industry dummies</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>adj. R^2</td>
<td>0.139</td>
<td>0.139</td>
<td>0.128</td>
<td>0.132</td>
</tr>
<tr>
<td>Observations</td>
<td>519</td>
<td>519</td>
<td>492</td>
<td>492</td>
</tr>
</tbody>
</table>

Robust t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Relationship between survey based training and expenditures and Kununu rating for training (firm-specific human capital) only borderline significant

Highly significant correlation between survey based marketing expenditures and Kununu rating for company image (brand equity)

**Source:** Training expenditures, marketing expenditures, turnover, number of employees and industry dummies from MIP 2017; Kununu data from company profiles with at least four ratings between January 2017 and August 2018
# OLS REGRESSIONS FACEBOOK

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>(1) ln(Image: Facebook likes)</th>
<th>(2) ln(Image: Facebook likes)</th>
<th>(3) ln(Image: Facebook likes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Marketing expenditures)</td>
<td>0.522*** (15.67)</td>
<td>0.454*** (8.23)</td>
<td>0.455*** (8.34)</td>
</tr>
<tr>
<td>ln(Turnover)</td>
<td>0.0589 (0.72)</td>
<td>0.106 (1.11)</td>
<td></td>
</tr>
<tr>
<td>ln(Number of employees)</td>
<td>0.0550 (0.65)</td>
<td>0.0178 (0.19)</td>
<td></td>
</tr>
<tr>
<td>Industry dummies</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>adj. R^2</td>
<td>0.261</td>
<td>0.265</td>
<td>0.364</td>
</tr>
<tr>
<td>Observations</td>
<td>944</td>
<td>944</td>
<td>944</td>
</tr>
</tbody>
</table>

Robust t statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Source: Marketing expenditure, turnover, number of employees and industry dummies from MIP 2017; Facebook likes from December 2017/January 2018

Highly significant correlation between survey based marketing expenditures and # of likes (brand equity)
MACHINE LEARNING

Idea: Internal expenditures could be estimated using (semi-) public data.

- Machine learning methods to predict internal firm expenditures with the help of pre-existing firm and platform data.
  - Internal expenditures based on platform data can be updated more often.

- Five machine learning regressions are trained for the analysis:
  - Neural networks (NN),
  - Random forests (RF),
  - K-nearest-neighbour (KNN) and
  - Support vector machines (SVM)
MACHINE LEARNING - FACEBOOK DATA

Predictive power for marketing expenditures based on Facebook data

<table>
<thead>
<tr>
<th>Features</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of employees</td>
<td></td>
</tr>
<tr>
<td>Turnover</td>
<td>Industry</td>
</tr>
<tr>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

The results are based on the mean values of five random train-test splits. The standard errors for MAE are mentioned in the brackets. A baseline model taking the non-transformed mean or median of the target variable in the train dataset as prediction has a MAE of 0.42 and 0.27. Data points within highest percentile of target variable were removed.

- The MIP and platform data can be used to a limited extent to estimate the internal expenditures of companies.
- MIP data alone is at most slightly better than the baseline.
- Adding our web data to MIP variables changes results only slightly.
- But we expect better results with more training data.
CONCLUSION/LIMITATIONS

- We are able to derive basic indicators for **brand equity** and **firm-specific human capital** from the platforms Facebook and Kununu.

- OLS regressions comparing platform indicators with firm-level survey data on marketing and training expenditure show a positive and significant relationship.
  - Various robustness checks confirm the validity of the results.

- Explorative machine learning approach shows that the platform data alone have little or no predictive power.

- Limited presence of smaller firms on online platforms.
  - Predominantly capturing medium-sized and larger firms.
  - Relatively low overlap between survey data and platform data.
Intangible Capital Indicators Based on Web Scraping of Social Media

Patrick Breithaupt (ZEW Mannheim) Reinhold Kesler (University of Zurich) Thomas Niebel (ZEW Mannheim) Christian Rammer (ZEW Mannheim)

Discussant: Charlotte Meng (ESCoE, NIESR)
11 November 2021
Summary

• This is an interesting and innovative research.

• The authors developed new indicators for intangible capitals (i.e. marketing and on-the-job training) using publicly available data from online platforms.

• OLS regressions show a positive and significant relationship between the newly-developed indicators and survey data.

• They also explored the possibility of predicting firm expenditures on brand equity and firm-specific human capital with machine learning methods.
#1: Representation of the sample

• Small firms are under-presented, as the authors have pointed out in section 3.3.1.

• Some industries are under-presented, e.g. hospitality are likely to invest more on social media marketing than manufacturing.

• While the authors acknowledged this as a disadvantage in the manuscript, it would be helpful to explore a bit more with the existing data. E.g. the distribution and summery stats of the 598 overlapping firms across size bands and industries;

• Elderly employees and customers are under-presented (The 2021 Social Media Users Demographics Guide | Khoros)

Age of internet users who use Facebook

• 86% of people ages 18-29 use Facebook
• 77% of people ages 30-49 use Facebook
• 51% of people ages 50-65 use Facebook
• 34% of people that are 65+ years old use Facebook
#2: Precision of the new measures

• Are Facebook ‘likes’ enough? Beyond ‘likes’, you can perhaps explore more dimensions of customer engagement including liking, sharing, commenting etc. (See, for example, Brodie, Ilic, Juric, and Hollebeek, 2013; Hollebeek, Glynn, and Brodie, 2014; Yost, Zhang and Qi, 2021)

• Kununu ratings for image to measure a firm’s marketing: when an employee rates a firm’s image, do they evaluate differently from customers? Customers view them outside the firms. But employees have personal experiences.

• MIP was conducted in 2017 and measures expenditures in that specific year. But data from online platforms are accumulative over several years.
Table 1. Social media use (May 2017 to May 2018).

<table>
<thead>
<tr>
<th>Social Media</th>
<th>Facebook</th>
<th>Twitter</th>
<th>Instagram</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of followers</td>
<td>160,875</td>
<td>2688</td>
<td>5141</td>
</tr>
<tr>
<td>Number of posts</td>
<td>323</td>
<td>241</td>
<td>108</td>
</tr>
<tr>
<td>Number of likes</td>
<td>7586</td>
<td>1766</td>
<td>15,399</td>
</tr>
<tr>
<td>Number of comments</td>
<td>648</td>
<td>n/a</td>
<td>767</td>
</tr>
<tr>
<td>Number of shares</td>
<td>1674</td>
<td>1039</td>
<td>n/a</td>
</tr>
<tr>
<td>Engagement total</td>
<td>9908</td>
<td>2805</td>
<td>16,166</td>
</tr>
<tr>
<td>Engagement rate</td>
<td>6.25%</td>
<td>104.4%</td>
<td>314.5%</td>
</tr>
</tbody>
</table>

*Engagement rate = the average number of interactions (i.e. comments, shares, likes) on number of followers during a selected period time, expressed as percentage.*

*Engagement total = the sum of number of interactions (i.e. comments, shares, and likes) during a selected period time.*

#3 Other suggestions

- How to utilise the indicators. Predictive power of those indicators with ML methods is not high, as is shown in section 6.

- It would be interesting to explore the data in a panel, if possible. That is, when a firm spend more on marketing/human capital in one year, social media engagement or rating increases or not in that year.

- The authors may provide more information about firm-specific human capital and brand equity, i.e. definitions and existing measures.

- Regardless of the innovative and exploratory essence of this research, I feel its contributions can be elaborated and highlighted in the introduction.