Valuing Data: Implications for Policy

ASSA Society of Government Economists
Session ‘Big Data: Value, National Accounts, and Public Policy’ 3 January 2020

Diane Coyle, Lawrence Kay, Stephanie Diepeveen, Jeni Tennison, Julia Wdowin

dc700@cam.ac.uk
A live policy issue

- Improved services, efficiency gains
  - high consumer valuations of digital services
  - new applications eg autonomous vehicles, IoT
- Potential of AI & big data for growth

BUT

- AI bias
- Privacy breaches & fines
- Commercial deals for public sector data
- Competition concerns ...
We have found that the profitability of both Google and Facebook has been well above any reasonable estimate of what we would expect in a competitive market for many years. In 2018 we estimated that the cost of capital for both Google and Facebook was around 9%, compared to actual returns on capital of over 40% for Google and around 50% for Facebook. This evidence is consistent with the exploitation of market power.

Sources: Furman 2019, CMA 2019
But above all.... sub-optimal provision & misallocation
Data economics

- Data is non-rival – technically a public good (club good, if access excludable) – and low marginal cost
Market mechanisms will not deliver socially optimal outcomes

- Private and public value diverge - +ve & -ve externalities
Too much or too little data both possible

- Value often due to combination with other data
‘Personal’ data has legal status but not a useful economic frame

- Data value unlikely to be related straightforwardly to volume (records, bits) because information content matters
Data has no ‘fundamental’ value
Two lenses on value of data

**Economic lens**
- Non-rival
- Externalities
- In-/Decreasing returns
- Option value
- High fixed/low marginal costs
- Complementary investments

**Information lens**
- Subject
- Generality
- Temporal characteristics
- Quality
- Sensitivity
- Interoperability & linkability
Existing approaches

- Stockmarket valuations
- Business outcomes – production or revenues
- Income generated in data value chain
- Cost-based methods
- Data markets
- Insurance premia
- Dark web markets
- Contingent valuation
Wide range of estimates for personal data

- 26 cents FT calculator based on data brokers
- £1-£200 Dark Web offers
- $2 FB average annual profit per active user
- $5/m WTP for privacy; willingness to forgo privacy $80/m (Winegar & Sunstein 2019)
- £0.005 ICO’s 2019 £500,000 fine on FB
- $125 FTC’s 2019 fine on Equifax
## Approaches to valuation?

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Issues</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diminishing/increasing marginal returns?</td>
<td>How granular is the necessary data? How much data is needed for prediction models?</td>
<td>Accuracy of predictive models</td>
</tr>
<tr>
<td></td>
<td>Is the holder using data accumulation as a source of market power?</td>
<td>Innovations and quality improvements in services</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Monopoly rents - profitability, absence of new entry</td>
</tr>
<tr>
<td>Externalities</td>
<td>Does aggregation, sharing/open data or joining different data sources add information?</td>
<td>Innovations and quality improvements in services</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Contingent valuation methods</td>
</tr>
<tr>
<td>Optionality</td>
<td>Does gathering more information provide scope for future process or quality improvements or innovation?</td>
<td>Real options methods?</td>
</tr>
<tr>
<td>Consequences</td>
<td>Are decisions made using the data highly consequential?</td>
<td>Value at Risk methods?</td>
</tr>
<tr>
<td>Costs</td>
<td>What costs need to be covered - data acquisition, cleaning, storage, skills/capabilities, governance</td>
<td>Harm to identified individuals (eg if later defrauded), loss of commercial confidentiality</td>
</tr>
<tr>
<td></td>
<td>What are the contingent costs - security breaches, loss of sensitive information, reputational damage, fines?</td>
<td>Risk assessments</td>
</tr>
</tbody>
</table>
Limitations of market approaches to valuation

- Market prices depend on policy framework – access rights, regulation etc.
- They do not fully capture forgone social welfare due to failure to recognise externalities
- To the extent they do internalise them, they are captured as monopoly rents
Data sharing & open data

The Data Spectrum

- **Internal access**: Employment contract + policies, Sales reports
- **Named access**: Explicitly assigned, Driving licences
- **Group-based access**: Via authentication, Medical research
- **Public access**: Licence that limits use, Twitter feed
- **Anyone**: Open licence, Bus timetable

Closed | Shared | Open

The Data Spectrum helps you understand the language of data. theodi.org/data-spectrum
Policy challenges

• Trade-off between incentivising/funding investment and shared access
  – Parallels with IP debates eg patent pools, compulsory licences?
• Regulatory framework will affect value of private assets
  – How should public bodies think about value of commercial deals?
• Institutional & regulatory design
  – Low trust
  – Asymmetric information, contract design under uncertainty, principal-agent problems
• Privacy/security concerns
• Current data sharing models have all required public sector initiative and/or funding
## Lessons from institutional economics: Ostrom’s governance principles

| There are clear boundaries and rules about who is entitled to what | Requires debate about ‘permissions’ - which entities can access which data? |
| Monitoring actions is feasible | Requires transparency about terms and conditions, and auditability |
| There are mechanisms for resolving conflicts | These could range from withdrawal of access permission up to legislated penalties |
| Individual responsibilities and benefits broadly balance | Requires transparency and better understanding of value exchanges that are occurring, including short term gain for long-term cost |
| Users themselves are responsible for monitoring and enforcement | A minimum requirement is transparency and contractual terms that enable monitoring and auditing of all subsequent data uses/transactions; may involve agents acting on behalf of data subjects |
| Sanctions for abuse are possible and graduated, getting progressively tougher | Suggests enhancing current approach - more enforcement |
| Decisions are legitimated by the participation of users | For individuals, opt outs need to be viable; importance of competition so users have alternatives; trustworthy institutions have representative governance bodies |
| Decisions are also legitimated by government recognition | Comprehensive data strategy and legal/regulatory framework will be needed |
Lessons from IO/regulatory economics?

- Patent pools, compulsory licences, network sharing → interoperability, open APIs, codes of conduct (but NB regulatory thickets)
- Incomplete contracts and asymmetric information → incentives to invest & share access to data
- Business models & charging structures → private/public mix
Thankyou!
dc700@cam.ac.uk