Innovation slowdown?

ARE WE OUT OF BIG IDEAS?

Dwindling gains in science, technology and medicine are a hidden drag on economic growth

BY GREG IP

By all appearances, we're in a golden age of innovation. Every month sees new advances in artificial intelligence, gene therapy, robotics and software apps. Research and development as a share of gross dom
time high
engineers
None o
ingful adv
of living.
Econor
panding v
such as ec
ings, then
more crea
"total fact

THE INNOVATION PARADOX

There is a yawning chasm between what innovation promises for the economy and what it is delivering. A Wall Street Journal series, running in full online,

provements in everyday life have been
cremental, not revolutionary. Houses, a
pliances and cars look much like they c
a generation ago. Airplanes fly no faste
than in the 1960s. None of the 20 most
prescribed drugs in the U.S. came to m
ket in the past decade.

Source: \

Centre for Economic Policy Research

POLICY INSIGHT No.63

September 2012

Is US economic growth over? Faltering innovation confronts the six headwinds

Robert J Gordon

Northwestern University and CEPR

1. Introduction

The prospects for future long-run US economic growth were already dismal in 2007 but were little noticed in the continuing euphoria over the invention of the internet and the related developments in information technology and communications (ICT). This Policy Paper pulls back from the past five years of financial crisis to pose a question with implications that will persist for decades even if the current international economic disorder is eventually resolved.

This paper is about US economic growth through 2007 and the future post-2007 path of potential or trend output for the subsequent 20 to 50 years. The analysis abstracts almost entirely from the
d since 2007.
questions of
ecovery have
output, and
end path and

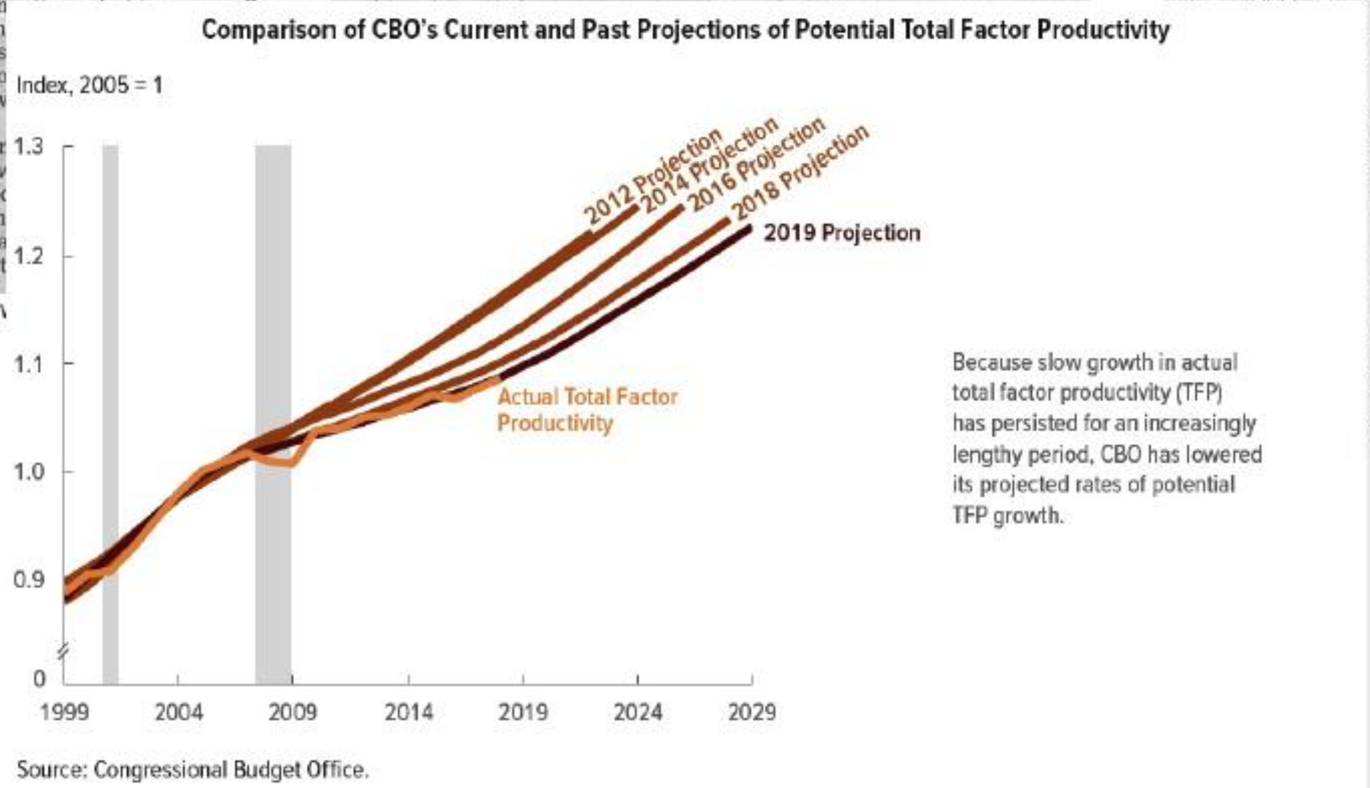
dox yet worth
n the context
e frontier of
ing have been
th century. If
r slows down,
r the slowing
ties for future

growth by all nations as the pace of productivity growth in the US fades out.

The paper makes these basic points:

1. Since Solow's seminal work in the 1950s, economic growth has been regarded as a continuous process that will persist forever. But there was virtually no economic growth before 1750, suggesting that the rapid progress made over the past 250 years could well be a unique episode in human history rather than a guarantee of endless future advance at the same rate.
2. The frontier established by the US for output per capita, and the UK before it, gradually began to grow more rapidly after 1750, reached its fastest growth rate in the middle of the 20th century, and has slowed down since. It is in the process of slowing down further.
3. A useful organising principle to understand the pace of growth since 1750 is the sequence of three industrial revolutions. The first (IR1) with its main inventions between 1750 and 1830 created steam engines, cotton spinning, and railroads. The second (IR2) was the most important, with its three central inventions of electricity, the internal combustion engine, and running water with indoor plumbing, in the relatively short interval of 1870 to 1900. Both the first two revolutions required about 100 years for their full effects to percolate through the economy. During the two decades 1950-70, the benefits of the IR2 were still transforming the economy, including air conditioning, home appliances, and the interstate highway system. After 1970, productivity growth slowed markedly, most plausibly because the main ideas of IR2 had by and large been implemented by then.

and other Policy Insights, visit www.cepr.org



Source: Congressional Budget Office.

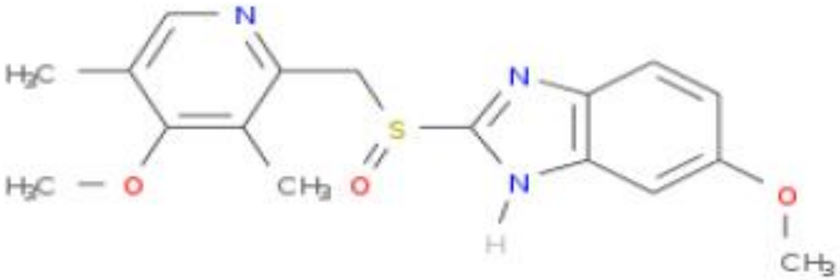
Quality and quantity of innovation?

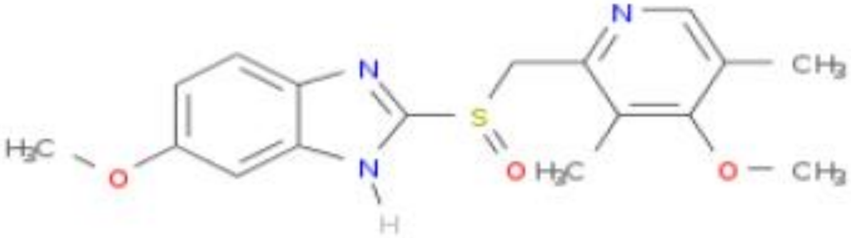
An innovation is valuable if

- It is novel (i.e., it does not copy an existing idea)
- It is useful (i.e., others use it or build on it)

We have a mechanism for tracking innovations: patents

Sadly, many patents are neither novel nor useful.





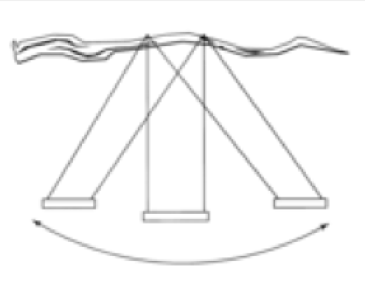
"A new medicine that isn't any better for ordinary heartburn than the one it will succeed" (WSJ, 2002)

Method of swinging on a swing

Abstract

A method of swing on a swing is disclosed, in which a user positioned on a standard swing suspended by two chains from a substantially horizontal tree branch induces side to side motion by pulling alternately on one chain and then the other.

Images (4)



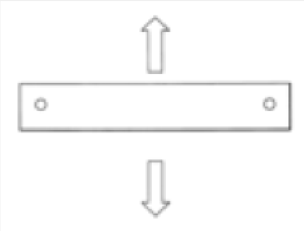


Figure 1

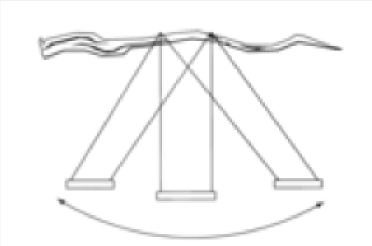
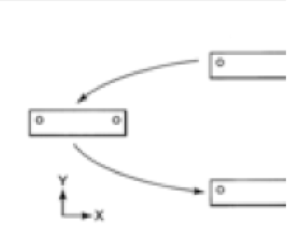


Figure 2



Figure

Classifications

A63G9/00 Swings

US6368227B1

US Grant

Download PDF Find Prior Art Similar

Inventor: Steven Olson

Original Assignee: Steven Olson

Priority date : 2000-11-17

Family: US (1)

Date	App/Pub Number	Status
2000-11-17	US09715198	Expired - Fee Related
2002-04-09	US6368227B1	Grant

Info: Patent citations (2), Cited by (11), Legal events, Similar documents, Priority and Related Applications

External links: USPTO, USPTO Assignment, Espacenet, Global Dossier, Discuss

Cited By (11)

Publication number	Priority date	Publication date	Assignee	Title
US6932710B1	2004-09-15	2005-08-23	William T. Hartin	Board swing
US20060036552A1 *	2003-01-31	2006-02-16	Microsoft Corporation	Secure machine counting
US20060293926A1 *	2003-02-18	2006-12-28	Khury Costandy K	Method and apparatus for reserve measurement
US20070232449A1 *	2004-11-26	2007-10-04	Nordisk Terapi As	Training apparatus
US20080293545A1 *	2004-11-26	2008-11-27	Redcord As	Training Apparatus
USRE41363E1 *	1995-11-21	2010-06-01	Samsung Electronics Co., Ltd.	Thin film transistor substrate
US20110003669A1 *	2004-11-26	2011-01-06	Redcord	Training apparatus
US20110239315A1 *	2009-01-12	2011-09-29	Ulla Bonas	Modular dna-binding domains and methods of use
US8420782B2	2009-01-12	2013-04-16	Ulla Bonas	Modular DNA-binding domains and methods of use
US8586526B2	2010-05-17	2013-11-19	Sangamo Biosciences, Inc.	DNA-binding proteins and uses thereof
US8697853B2	2009-12-10	2014-04-15	Regents Of The University Of Minnesota	TAL effector-mediated DNA modification
Family To Family Citations				

* Cited by examiner, † Cited by third party, ‡ Family to family citation

Automated patent quality measurement

Need to ‘read’ 9MM USPTO patent since 1836 to see how each is

- Novel: distinct from previous patents
- Impactful: related to subsequent patents

Automate this:

- Featurize patent text to capture info that was key *when it appeared*
- Train ML to infer past and future patents from these features

Low past similarity implies novelty, good future predictors have impact.

Automated patent quality measurement

Need to ‘read’ 9MM USPTO patent since 1836 to see how each is

- Novel: distinct from previous patents
- Impactful: related to subsequent patents

Document featurization: words *w* in patent *d* at time *t* are given value

$$TFBIDF_{wd} = \frac{count_{wd}}{\sum_{w'} count_{w'd}} \times \log \frac{D_t}{1 + D_{tw}}$$

where D_{tw} is the # patents by t containing w and $D_t = \sum_w D_{tw}$

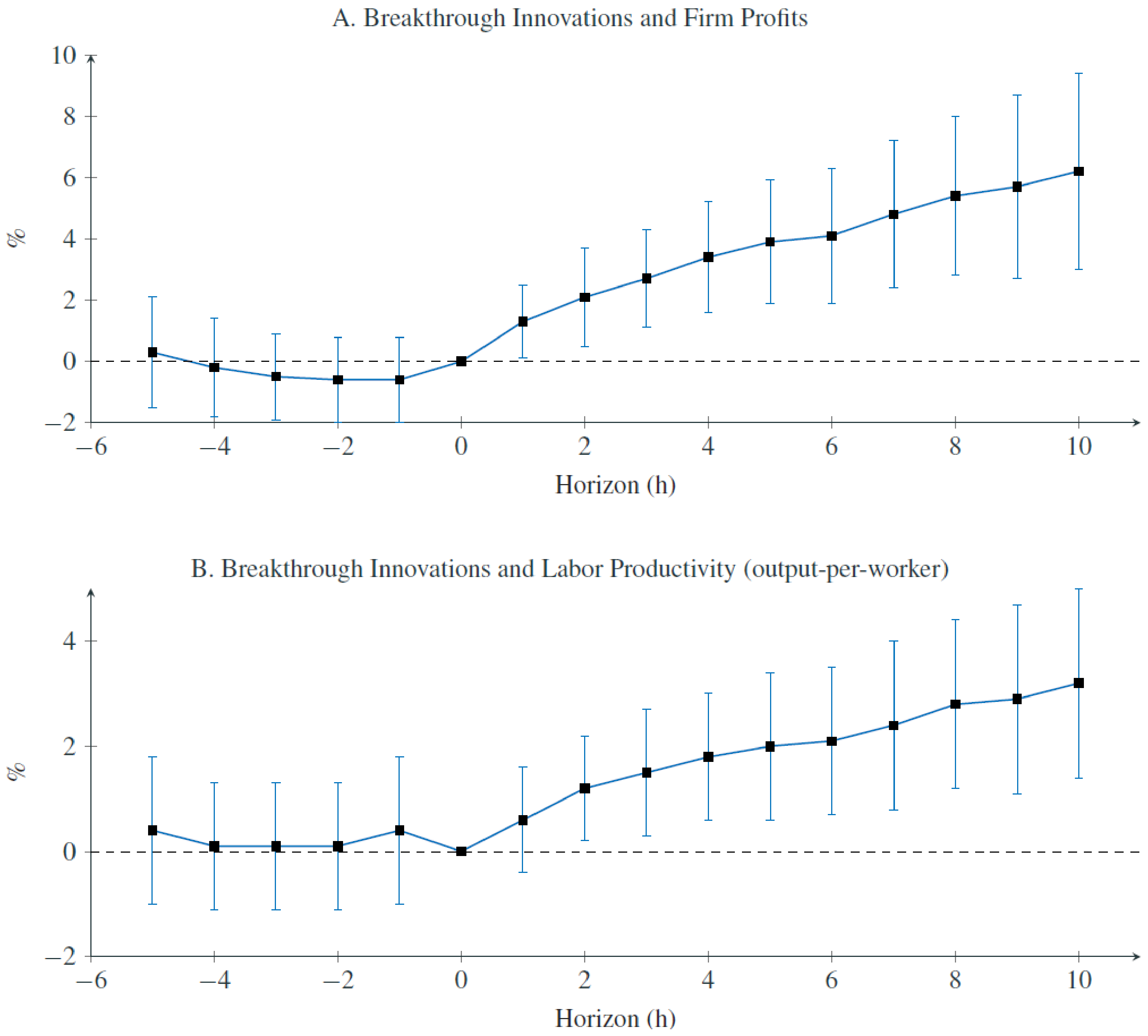
Automated patent quality measurement

Need to 'read' 9MM USPTO patent since 1836 to see how each is

- Novel: distinct from previous patents
- Impactful: related to subsequent patents

Connecting to past and future:

- Calculate cosine similarity on TFBIDF vectors, ρ_{dk} for patents d, k
- Sum similarities for each patent with future (f_d) and past patents (b_d)
- Define *quality* as ratio of future over past similarity $q_d = \log \frac{f_d}{q_d}$


A. Breakthrough Innovations and Firm Profits
%
Horizon (h)
B. Breakthrough Innovations and Labor Productivity (output-per-worker)
%
Horizon (h)

An index of innovation

Decompose quality for patent d from year t as $q_d = \alpha_t + \beta_d$

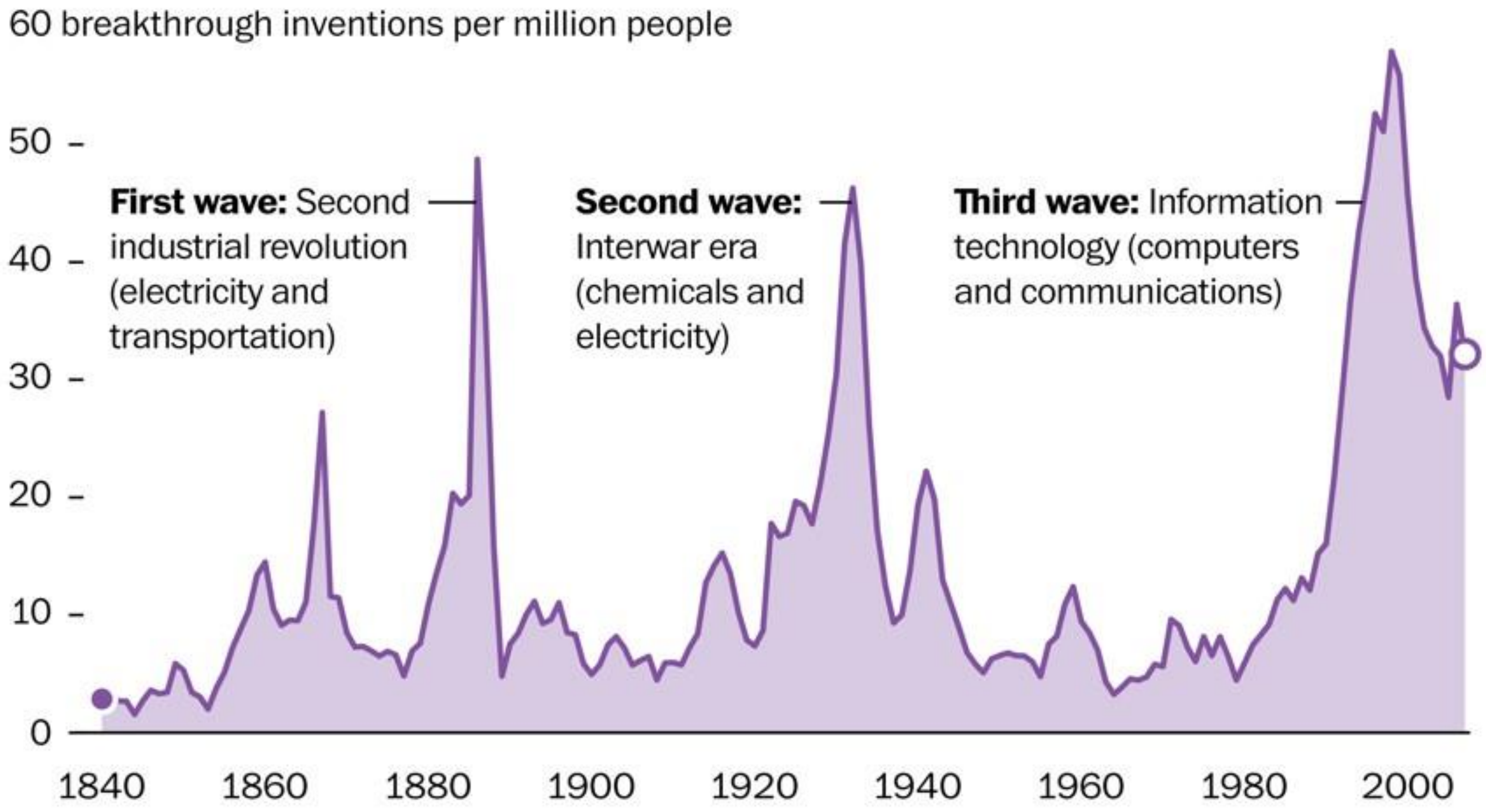
(control for the temporal changes in language usage)

Count the number of patents for year t with β_d in the 95$^{\text{th}}$ percentile

(dampen estimation noise)

When the most influential U.S. inventions were patented

Annual count of breathrough patents, adjusted for population

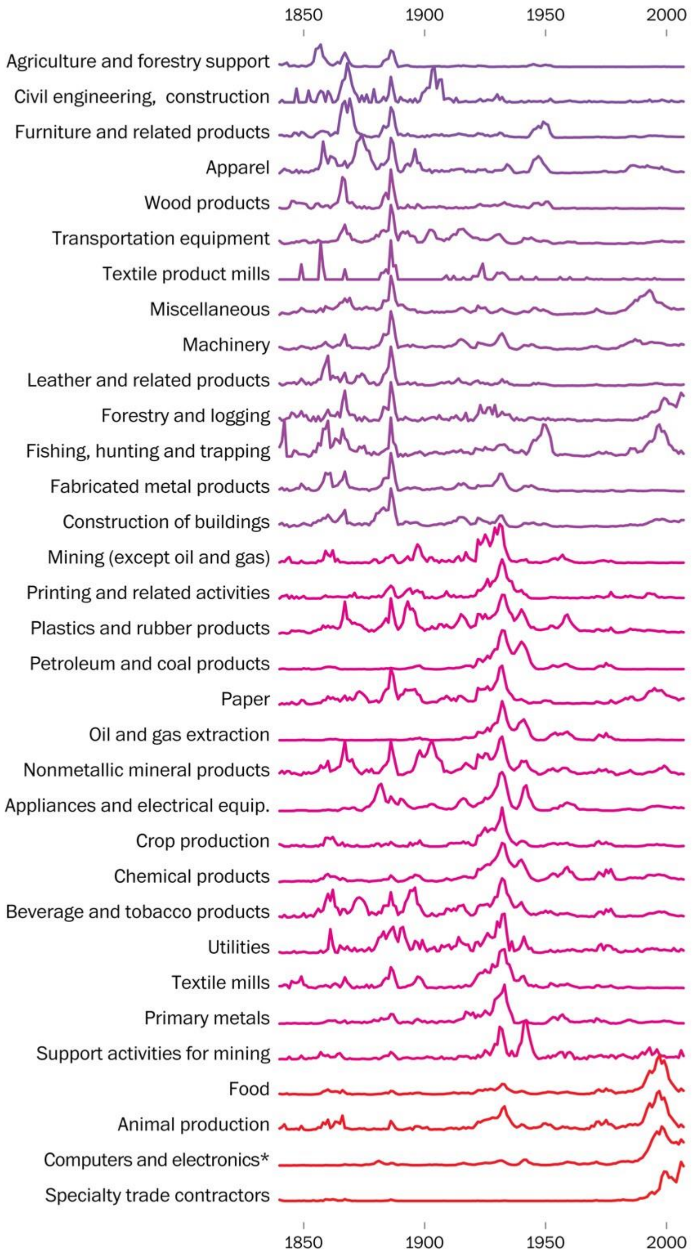


Source: Analysis of U.S. Patent Office data by Bryan Kelly, Dimitris Papanikolaou, Amit Seru, Matt Taddy

THE WASHINGTON POST

When each industry's top innovations were patented

Indexed relative to each industry's individual peak

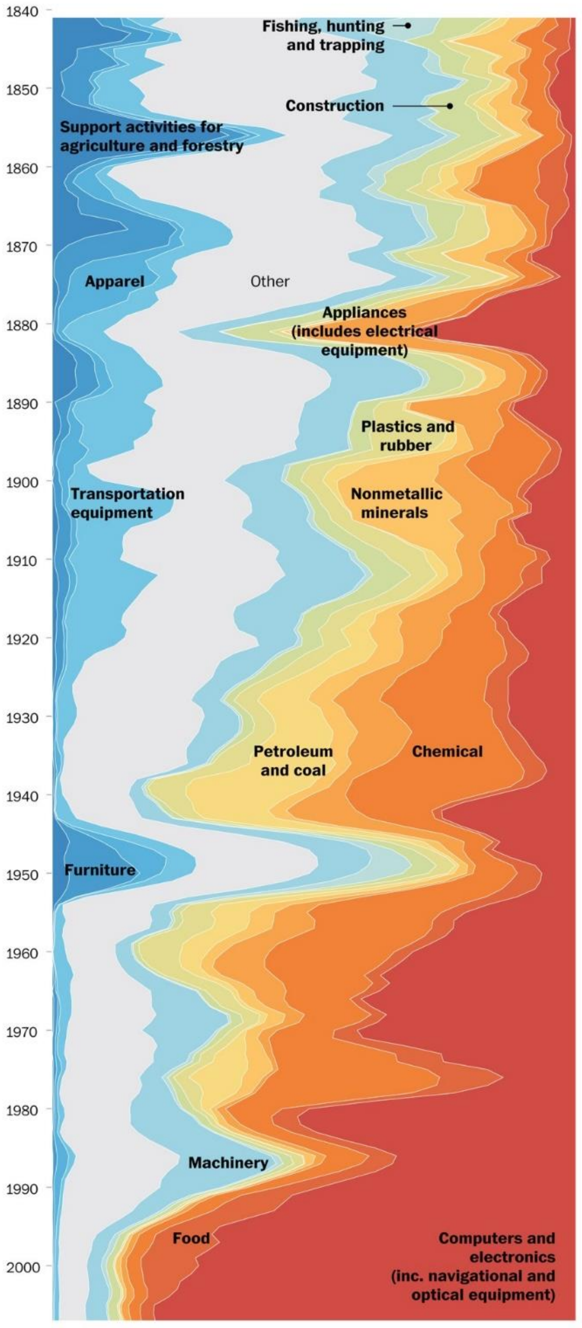


* Includes navigational and optical equipment

Source: Analysis of U.S. Patent Office data by Bryan Kelly, Dimitris Papanikolaou, Amit Seru, Matt Taddy

THE WASHINGTON POST

Distribution of influential U.S. inventions by industry



Source: Analysis of U.S. Patent Office data by Bryan Kelly, Dimitris Papanikolaou, Amit Seru, Matt Taddy

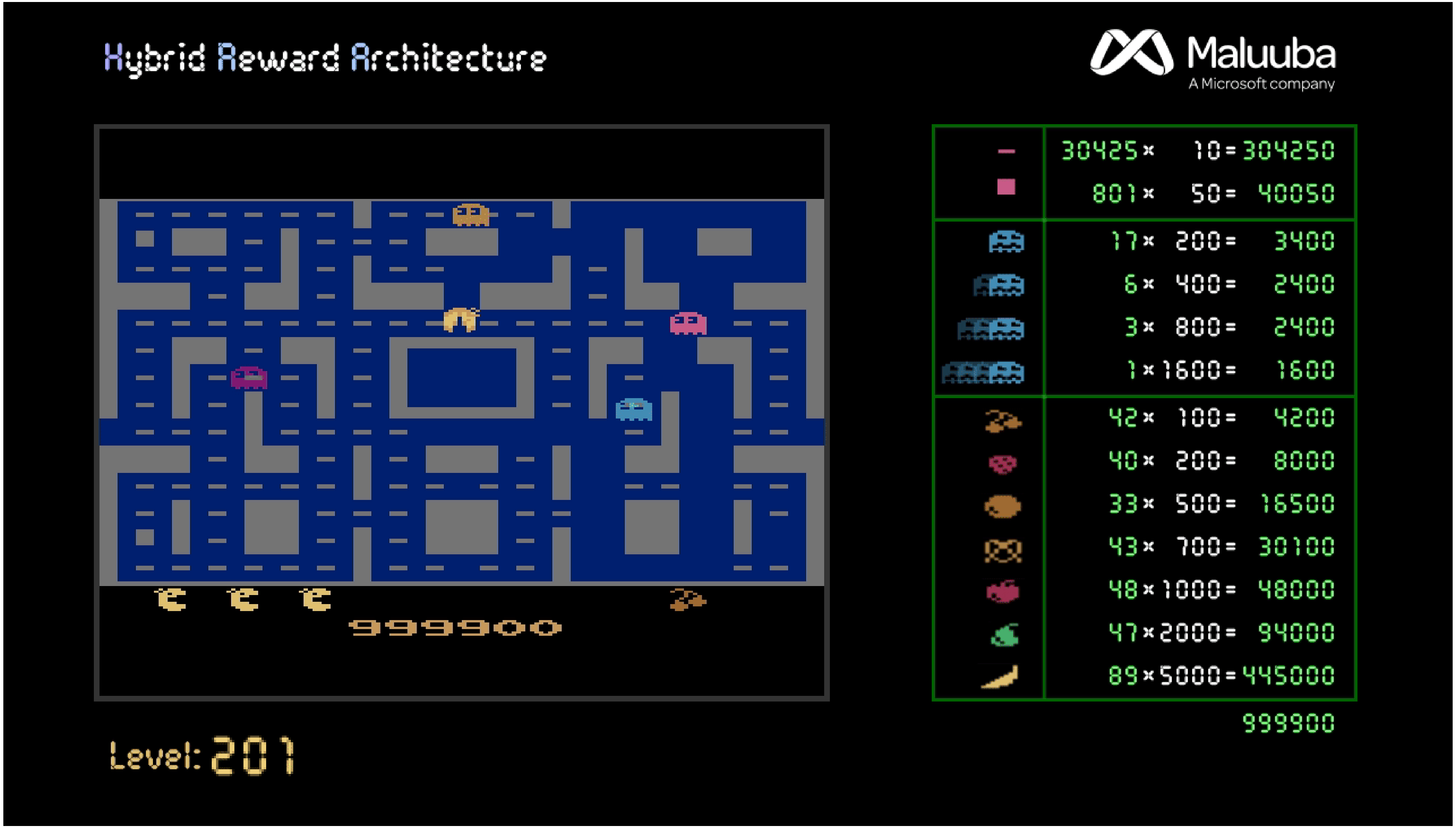
THE WASHINGTON POST

What is AI?

Domain Structure + Data Generation + General Purpose ML

Domain Structure	Data Generation	General Purpose ML
Econ/Biz Framework Hands off the Wheel Causal Inference	Reinforcement Learning Sensor/Camera networks Simulation/GANs	Auto ML, Sagemaker, Inferentia, DNNS, GPUs, Comp Vision, NLU, NLG

Self-training structures of ML predictors that automate and accelerate human tasks



https://www.youtube.com/watch?time_continue=13&v=zQyWMHFjewU

Econometrics breaks systemic questions into sets of prediction tasks

- Prediction after controlling for confounders
- Two-stage prediction with IVs
- Heterogeneous treatment effect prediction
- Structural equation systems

Machine Learning can automate and accelerate
tasks in applied econometric workflows

Example: short-term price elasticity

If I drop price 1%, by what % will quantity sold increase?

Problem: both prices and sales respond to underlying demand

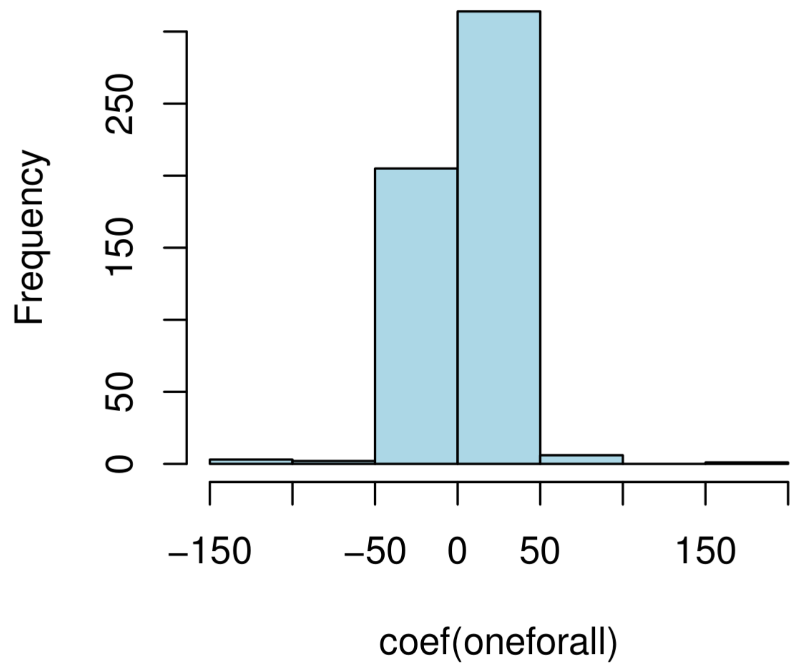
Need a causal effect of price on sales, not their co-movement

Beer Data

A single shared elasticity gives tiny -0.23

Separate elasticity for each: noisy zeros

We need to group the products together using brand, pack, etc.



Beer Elasticity

Say $w_{bk} = 1$ if word k is in description for beer b

@transaction t: $y_{tb} = \gamma_b \, p_{tb} + f_t(\boldsymbol{w}_b) + \varepsilon_{tb}$,

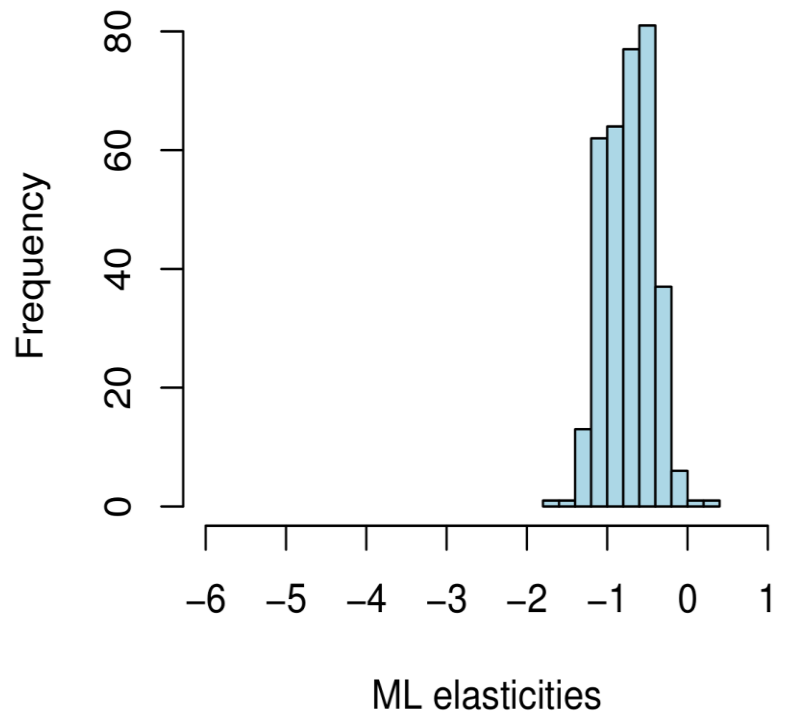
$$\gamma_b = \boldsymbol{w}_b' \boldsymbol{\beta}$$

$$p_{tb} = h_t(\boldsymbol{w}_b) + \nu_{tb}$$

Creates a large number of parameters

Just throw it all in a lasso?

Yields unbelievably small elasticities



The naïve ML conflates two problems:
selecting controls and predicting response after controlling for confounders.

Instead, use Orthogonal ML *(Chernozhukov et al, 2016 and earlier)*

- Estimate nuisance functions $\mathbb{E}[y_{tb}|\boldsymbol{w}_b]$ and $\mathbb{E}[p_{tb}|\boldsymbol{w}_b]$
- Orthogonalize the score against these nuisance functions (data split)
- Then estimation for $\boldsymbol{\gamma}$ is robust to slow-learned nuisances

Estimation breaks into a series of ML tasks:

1. Predict sales from the demand variables: $y_{tb} \approx g(t, \boldsymbol{w}_b)$
2. Predict prices from the demand variables: $p_{tb} \approx h(t, \boldsymbol{w}_b)$
3. Get OOS residuals: $\tilde{\boldsymbol{y}}_t = \boldsymbol{y}_t - \hat{g}_{\bar{t}}(t, \boldsymbol{w}_b)$, $\tilde{\boldsymbol{p}}_t = \boldsymbol{p}_t - \hat{h}_{\bar{t}}(t, \boldsymbol{w}_b)$
4. And fit the final regression: $\mathbb{E}[\tilde{\boldsymbol{y}}_t] = \boldsymbol{\Gamma}\, \tilde{\boldsymbol{p}}_t = \mathrm{diag}(\boldsymbol{\gamma})\, \tilde{\boldsymbol{p}}_t$

Orthogonal ML for Beer

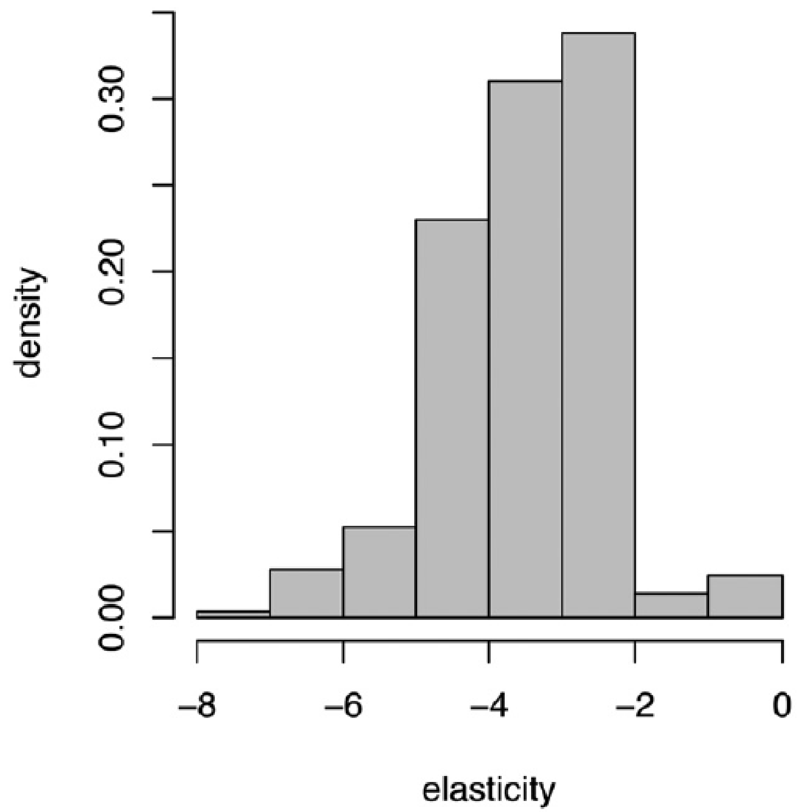


FIGURE 6.3: Beer-specific elasticities obtained by lasso regression on the orthogonal ML residuals.

The text encodes a natural hierarchy

Many beers are `IPA` or `Cider`

But individual brands also load

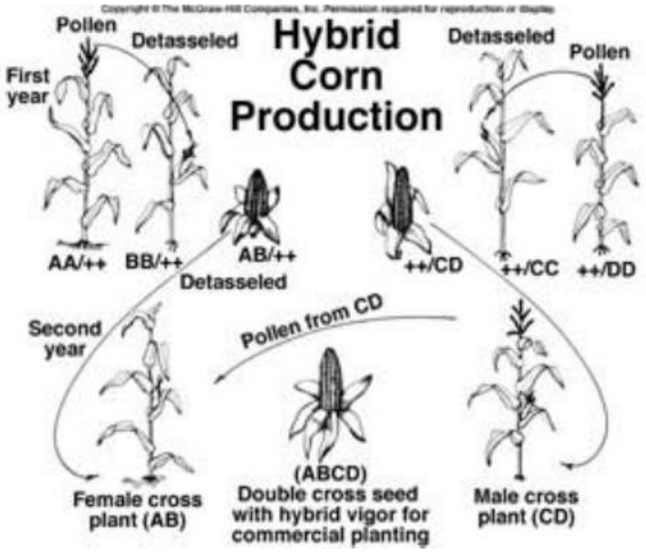
Most Price Sensitive

```
> names(sort(el)[1:5])
[1] "GUINNESSS DRAUGHT 6PK BTL
[2] "GUINNESS DRAUGHT 4PK CAN
[3] "PYRAMID OUTBURST IMP IPA 6PK
[4] "ELYSIAN IMPORTAL IPA 6PK
[5] "PYRAMID OUTBURST IMP IPA 12PK
```

Least Price Sensitive

```
> names(sort(-el)[1:5])
[1] "2 TOWNS CRISP APPLE CIDER
[2] "2 TOWNS BAD APPLE CIDER
[3] "ATLAS BLKBRY APPLE CIDER
[4] "D'S WICKED BAKED APPLE CIDER
[5] "D'S WICKED GREEN APPLE CIDER
```

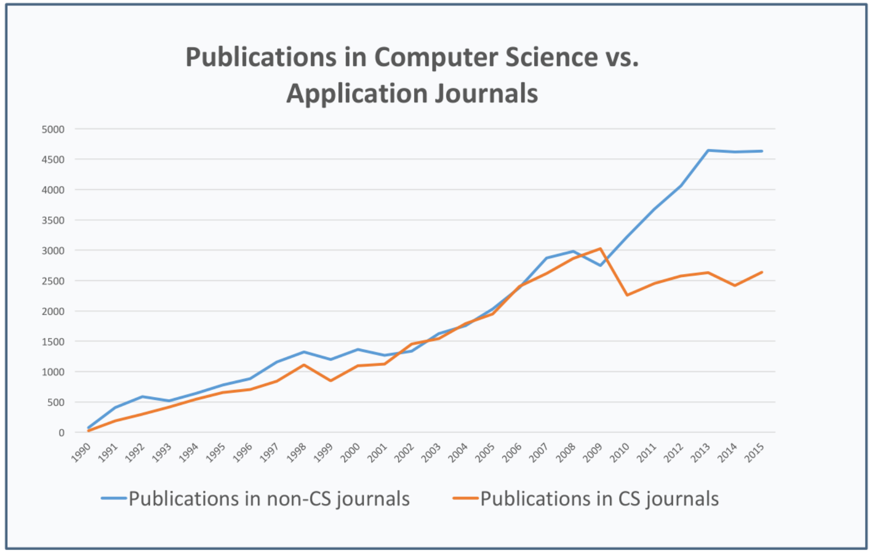
The Economics *of* AI



DNNs are GPT and 'method for invention'

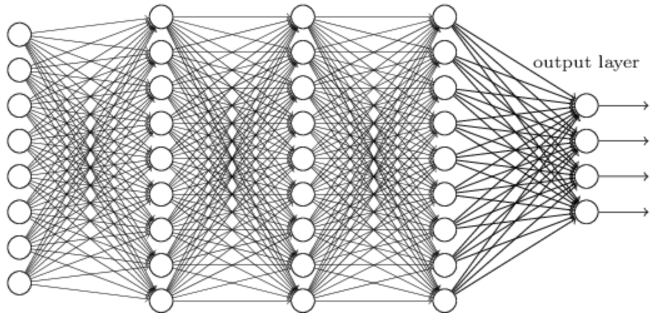
- Broad impact, up and down the value chain
- Gets better, faster, and cheaper in time
- Can suffer from underinvestment
- Productivity gains lag invention

AI research in computer science journals vs. other application sectors.



Graph from Cockburn/Henderson/Stern

DEEP NEURAL NETWORKS



Train faster, generalize better:
Stability of stochastic gradient descent

Adaptive Subgradient Methods for Online Learning and Stochastic Optimization*

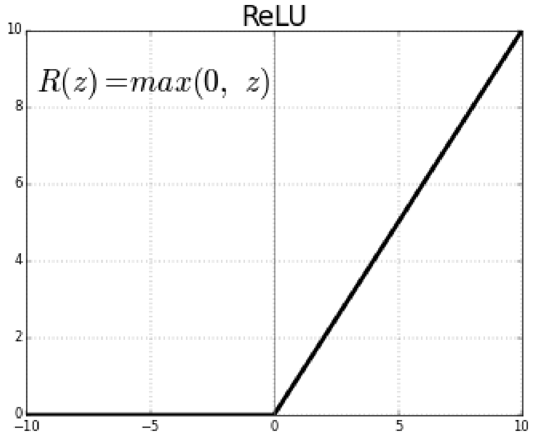
$$\mathcal{L} + \beta \underbrace{\frac{1}{2}\|w\|_2^2}$$

Dropout: A Simple Way to Prevent Neural Networks from Overfitting

TensorFlow

mxnet

PyTorch

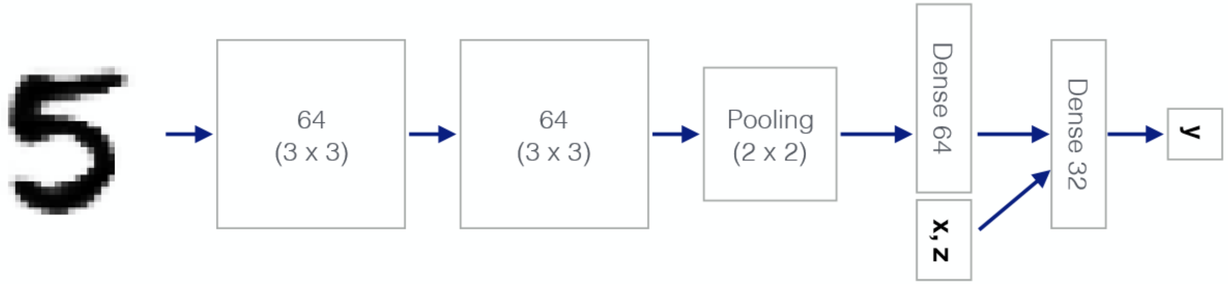


K

GLUON

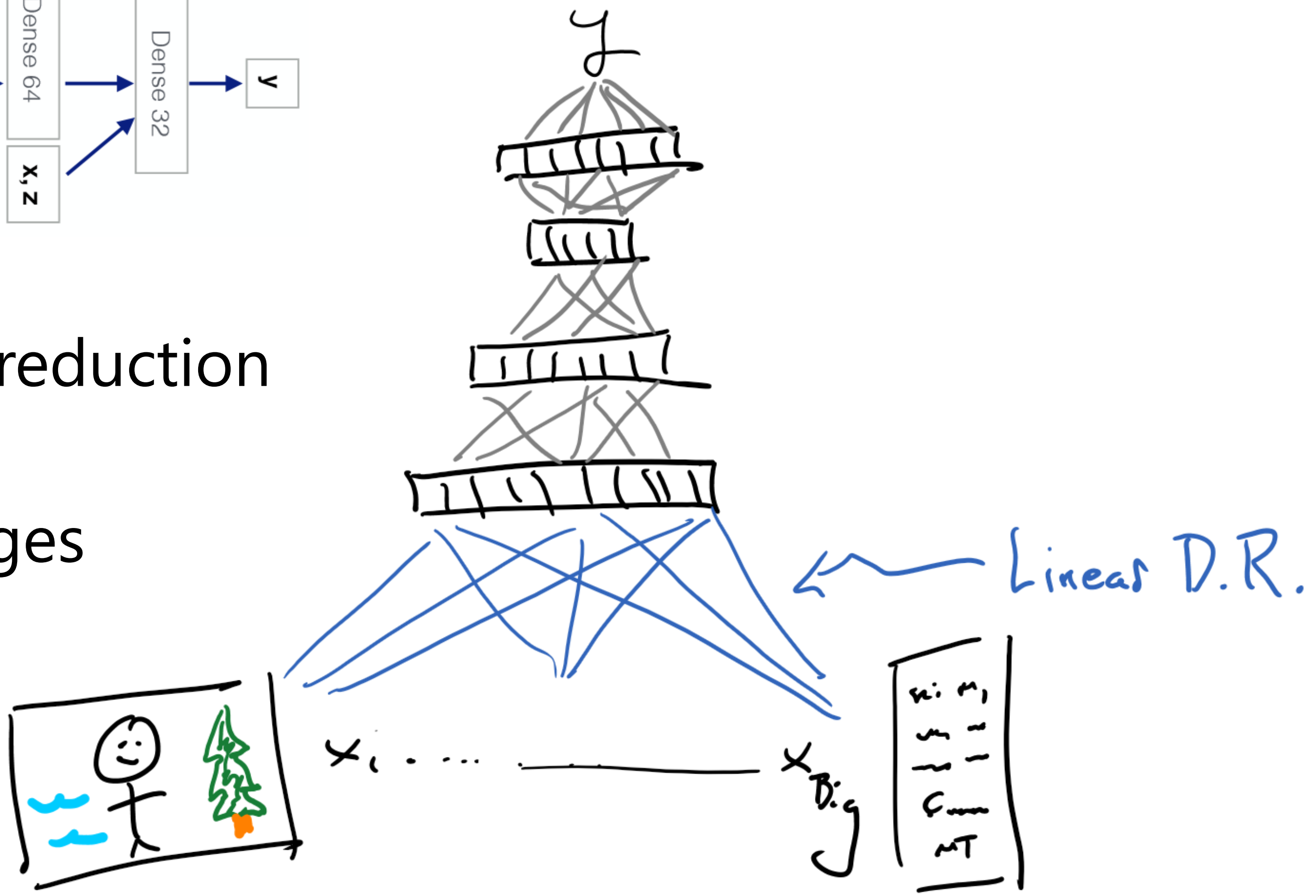
HOROVOD

Deep nets are not nonparametric sieves



The 1st layer is a big dimension reduction

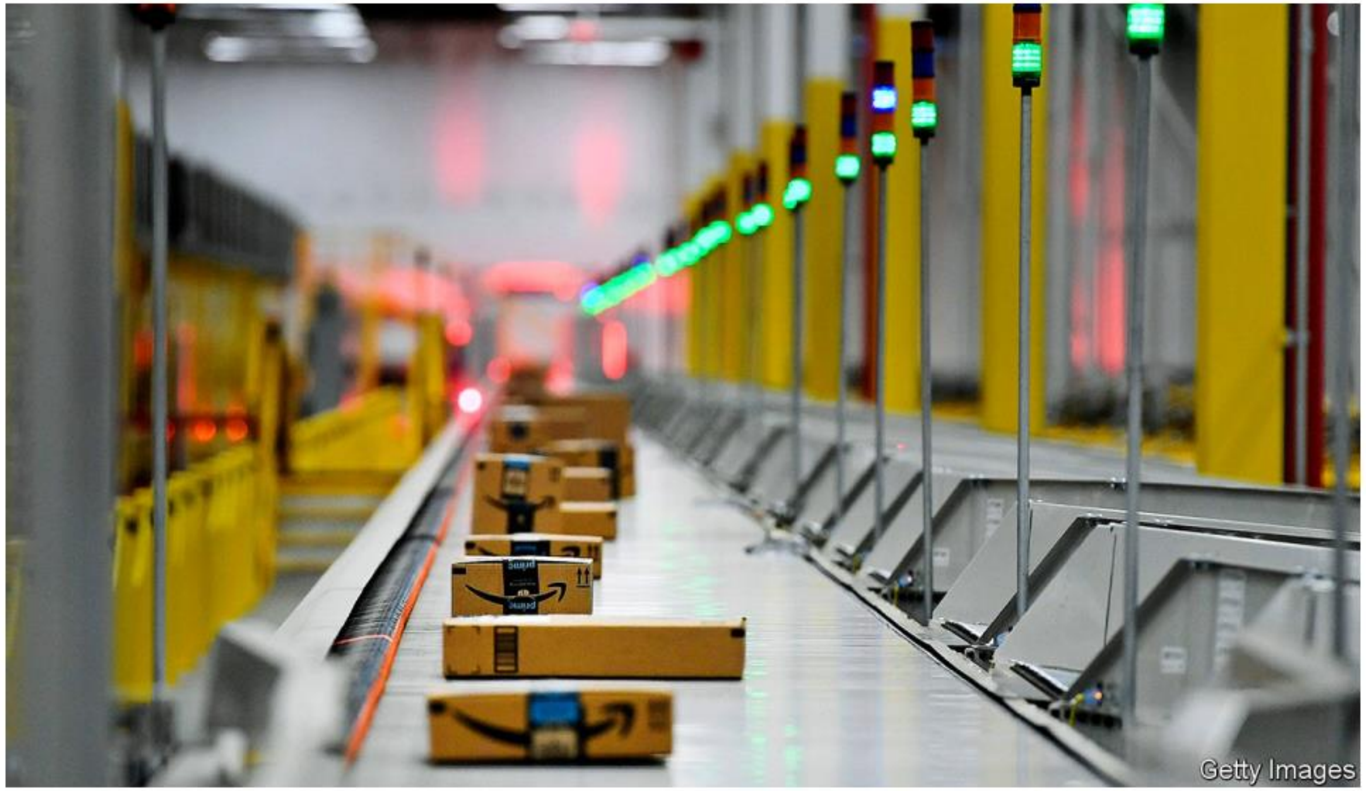
- word embedding for text
- matrix convolution for images



The learning machine

Amazon’s empire rests on its low-key approach to AI

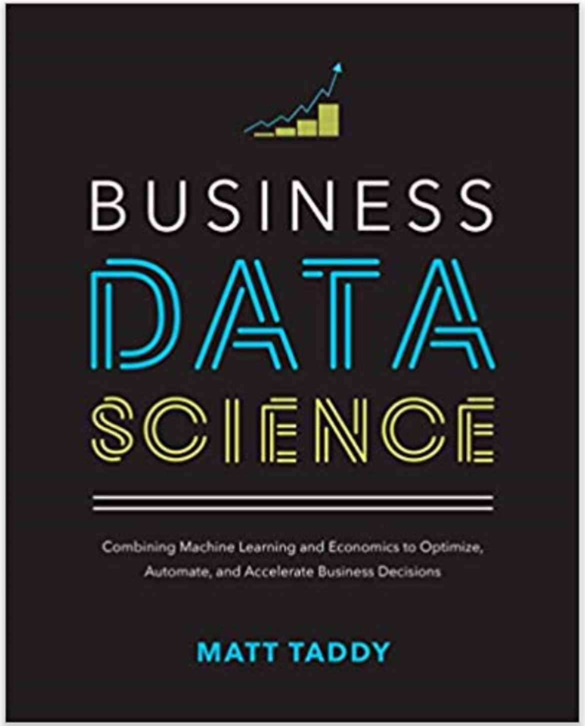
Unflashy but high-powered machine learning powers everything from its fulfilment centres to the cloud



- Key business reviews ask: *How are you using ML?*
- Science and Business break big questions into ML tasks
- Need to answer “Why?”
- AI impacting physical world – boxes, robots, humans!
- Improving ML: Sagemaker, Auto ML, Inferential, CV …

Business Data Science: Combining Machine Learning and Economics to Optimize, Automate, and Accelerate Business Decisions 1st Edition

by Matt Taddy (Author)



ISBN-13: 978-1260452778

eTextbook	Hardcover	Other Sellers
$24.20	$25.47	See all 2 versions

Pre-order

prime $25.47

This title will be released on August 23, 2019.

Ships from and sold by Amazon.com.

List Price: ~~$39.00~~

Save: $13.53 (35%) Pre-order Price Guarantee.

1 New from $25.47

Deliver to Matthew - Seattle 98112

Qty: 1

Pre-order: Add to Cart

Pre-order now

Q&A