Engineering Value: The Returns to Technological Talent and Investments in Artificial Intelligence

Daniel Rock (UPenn - Wharton)
@danielrock

AEA 2021

Advice for workers that works for firms too

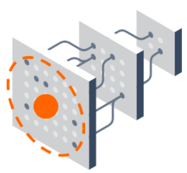
“If you are looking for a career where your services will be in high demand, you should find something where you provide a scarce, complementary service to something that is getting ubiquitous and cheap. So what’s getting ubiquitous and cheap? Data. And what is complementary to data? Analysis.” – Hal Varian, 2008 (Freakonomics Blog)

A specific story about TensorFlow, but with a general lesson

Why TensorFlow

TensorFlow is an end-to-end open source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML powered applications.

About →



Easy model building

Build and train ML models easily using intuitive high-level APIs like Keras with eager execution, which makes for immediate model iteration and easy debugging.

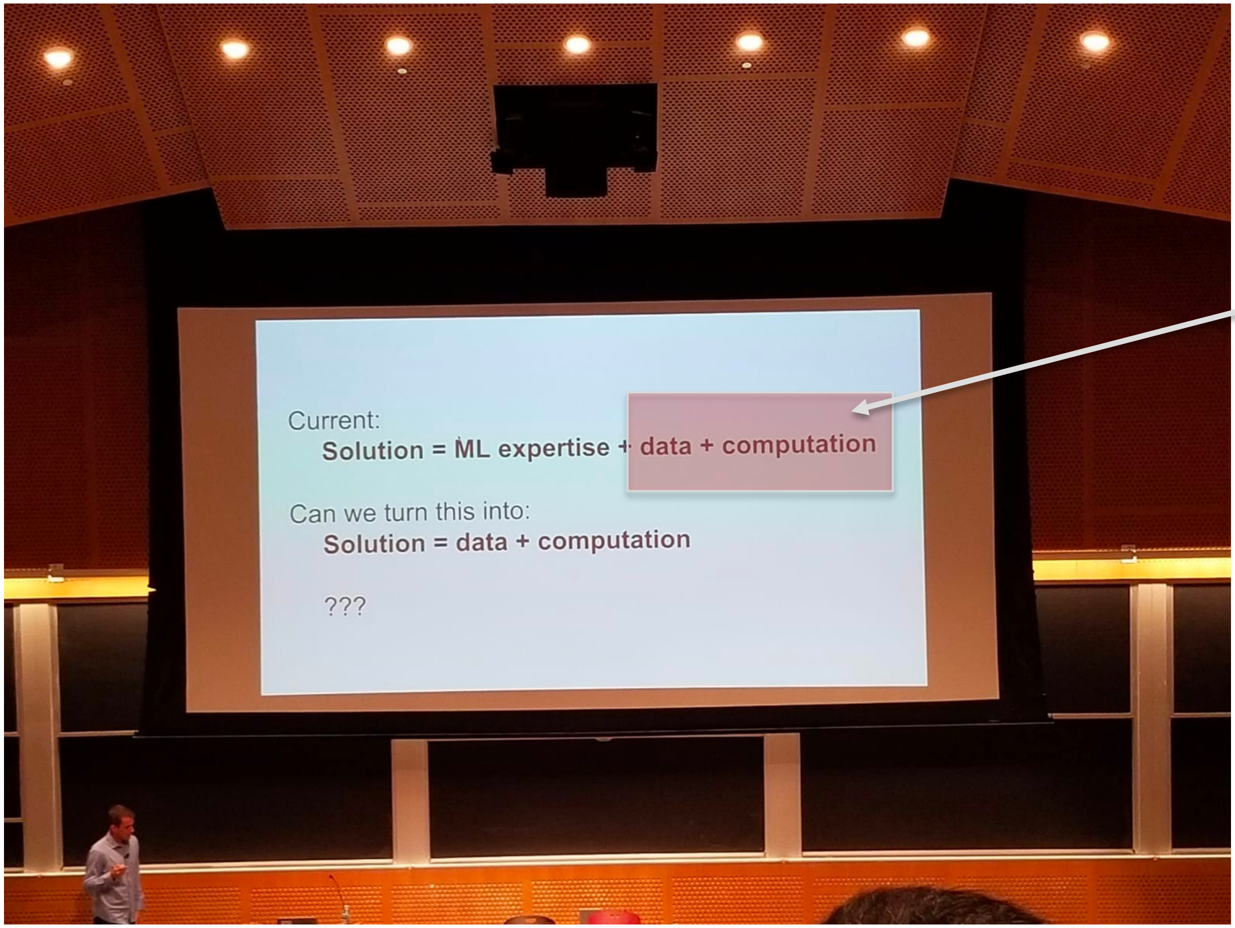
Robust ML production anywhere

Easily train and deploy models in the cloud, on-prem, in the browser, or on-device no matter what language you use.

Powerful experimentation for research

A simple and flexible architecture to take new ideas from concept to code, to state-of-the-art models, and to publication faster.

TensorFlow affects talent



Google has lots of assets that do this

Research Question: Who benefits from investments in technological talent?

1) Theory: How firm-specific intangible and human capital helps capture rents from high-skilled employees
   - …even after wage premia are paid
   - Price effects on sunk investments a likely story
   - Other possibilities:
     - Contemporaneous productivity
     - Wage declines
     - Capital quantity increases

2) Detailed LinkedIn panel data on job history and *skills* in engineering and technology

Research Question: How do firms benefit from investments in technological talent?

3) A natural experiment: Which firms capture value from a technology shock that makes skills abundant?

- Google's TensorFlow dramatically decreased the cost of learning to do *deep learning*
- Firms whose engineers possessed more AI skills captured more value from the TensorFlow launch → 4-7% market value increase!
  - Middle quintile firms get the biggest boost!
  - Results hold up with alternative specifications

4) Does hiring more engineers increase firm value?

- Not when talent is available! But *average* value is ~$850k per engineer

Technological Labor is a driver of productivity

- …and market value (Jaffe 1986; B.H. Hall et al. 2005; R.E. Hall 1993, 2006; Tambe and Hitt 2012; Tambe 2014)
- Often related to intangible assets and co-invention (Bresnahan et al. 1996; Hall 2001; Bresnahan et al. 2002; Greenstein and Nagle 2014; Saunders and Tambe 2015; Peters and Taylor 2017; Brynjolfsson, Rock, and Syverson 2018; 2019)
- These are (often) firm-specific!

The employer bargains away some of the surplus for firm-specific tasks

- *Non-monetary compensation*: *Working with cutting-edge technology* (Stern 2004; Roach and Sauermann 2010; Mas and Pallais 2017)
- *Monopsony Power coming from frictions* (Bhaskar et al. 2002; Ashenfelter et al. 2010; Azar et al. 2018; Stole and Zwiebel 1996a,b)
- *Firm-specific Capital* (Brynjolfsson et al. 2002; Coff and Raffiee 2015; Eisfeldt and Papanikolaou 2014; Kline et al. 2018)
- *Mobility* (Campbell et al. 2012; Benmelech et al. 2018)

We're at an important crossroads with respect to technology workers

- *STEM workers* (Peri et al. 2015; Kerr et al. 2015; Ding et al. 2017; Glennon 2018)
- *Unexplained productivity differences* (Syverson 2011; Andrews et al. 2015; R.E. Hall 2018)
- *The tools are changing* (Teodoridis 2017; Thompson 2017; Ewens, Nanda, and Rhodes-Kropf 2018; Zyontz 2018; Agrawal et al. 2018, Choudhury et al. 2018)
- *And breakthroughs are harder to find?* (Jones 2009; Cockburn et al. 2018; Webb et al. 2018)

Value a firm by valuing its assets

$$Firm\ Value\ Solution\text{:}\ \ V(0) = \sum_{j=1}^{J} \lambda_j(0) K_j(0)$$

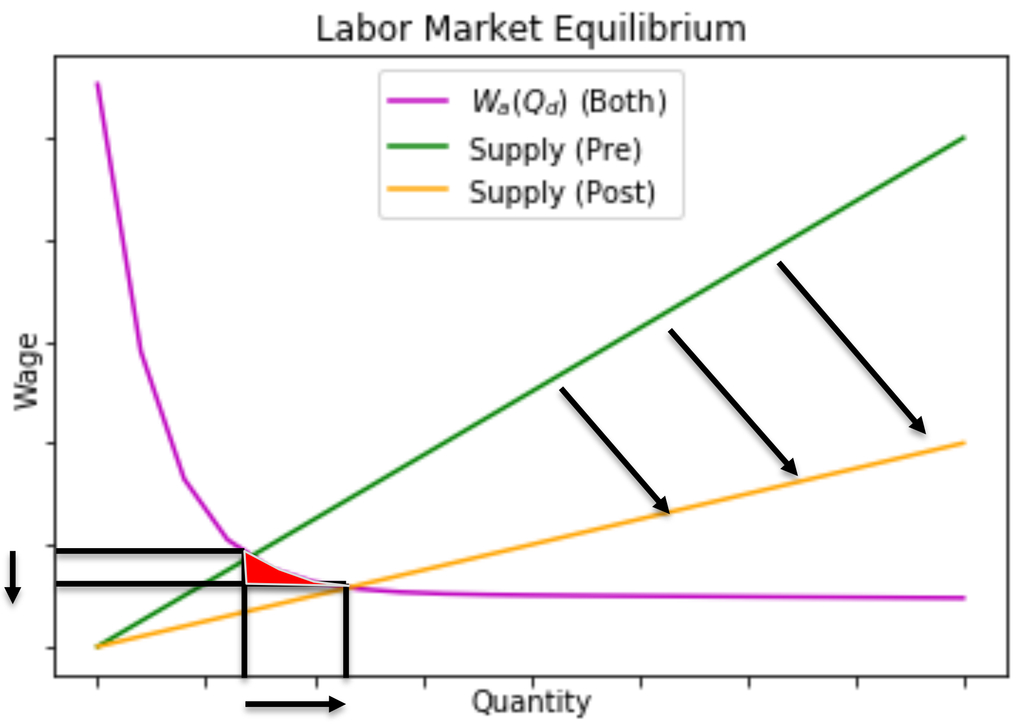
$$\int_0^\infty \left[\boxed{\sum_{j=1}^{J} \left(pF_{K_j} K_j + pF_{I_j} I - z_j I_j \right)} + \boxed{\sum_{i=1}^{L} \left(pF_{L_i} L_i - w_i \right)} \right] u(t)dt = V(0) = V_K + V_L$$

Capital and investment Value

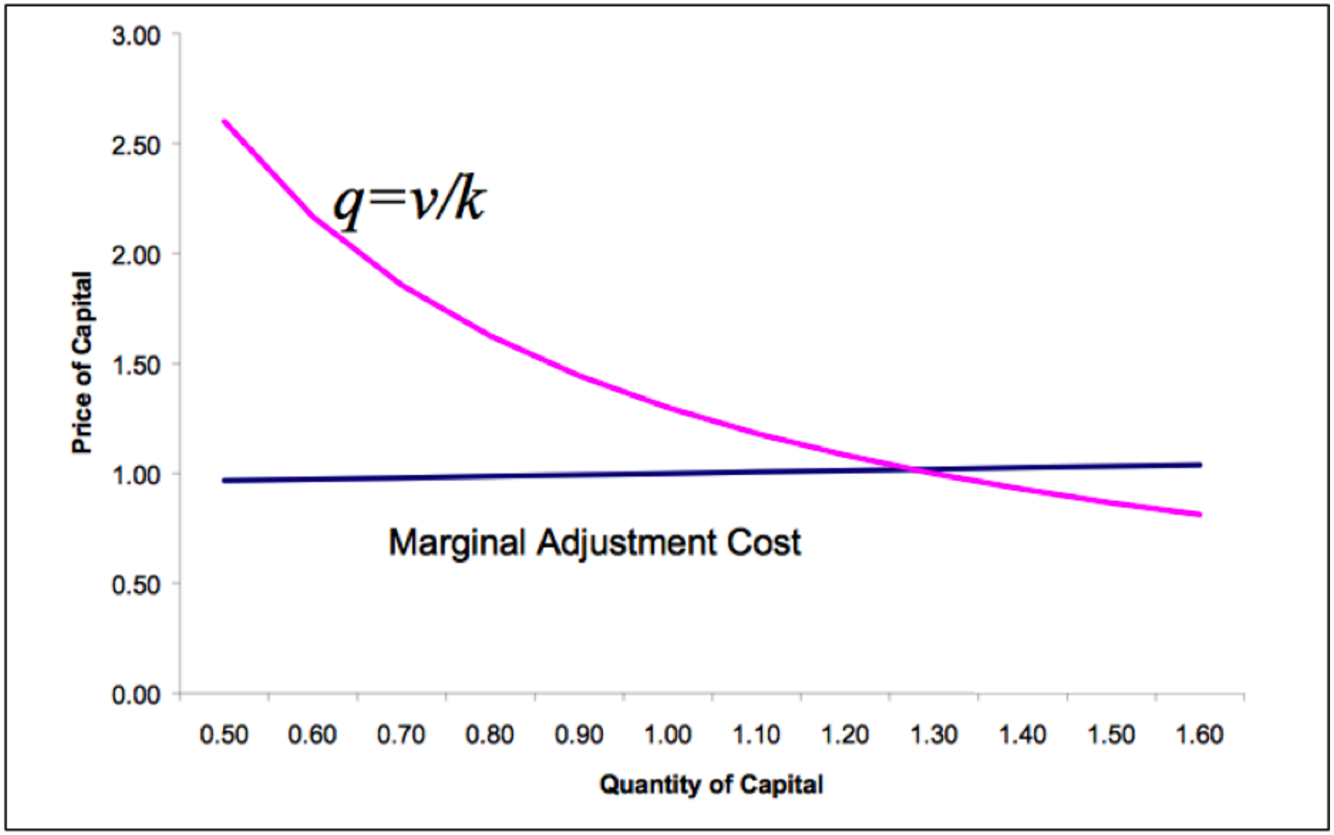
Labor Value
(Often assumed to be zero. It's probably not zero!)

Setup Value Wedge S.I.D.

Returns to technological skill shocks can be shared by the employer – Does this happen with TensorFlow?

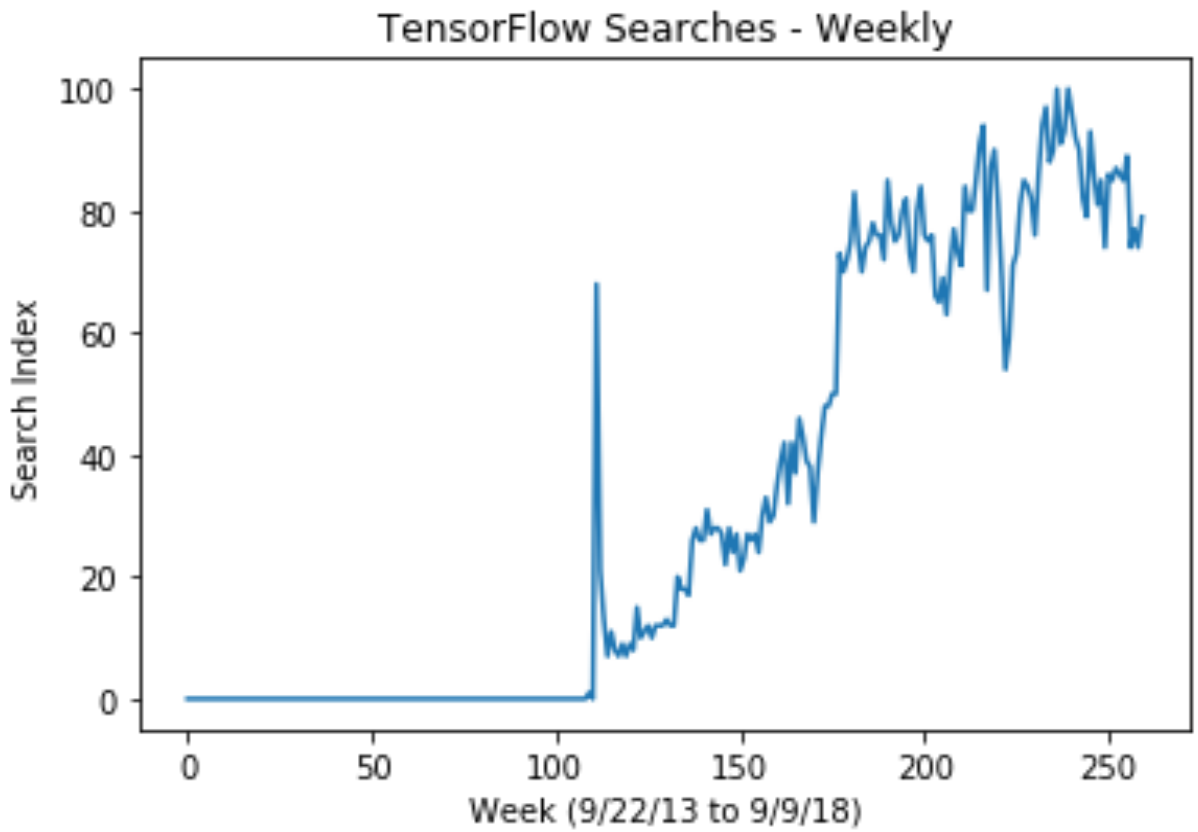


Asset price equilibrium is partly driven by the prices of complements



Source: Hall (2001)

TensorFlow was a surprise



Source: Google Trends

TensorFlow is (relatively) easy to learn

```python
import tensorflow as tf
mnist = tf.keras.datasets.mnist

(x_train, y_train),(x_test, y_test) = mnist.load_data()
x_train, x_test = x_train / 255.0, x_test / 255.0

model = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(),
  tf.keras.layers.Dense(512, activation=tf.nn.relu),
  tf.keras.layers.Dropout(0.2),
  tf.keras.layers.Dense(10, activation=tf.nn.softmax)
])
model.compile(optimizer='adam',
              loss='sparse_categorical_crossentropy',
              metrics=['accuracy'])

model.fit(x_train, y_train, epochs=5)
model.evaluate(x_test, y_test)
```

Without TF

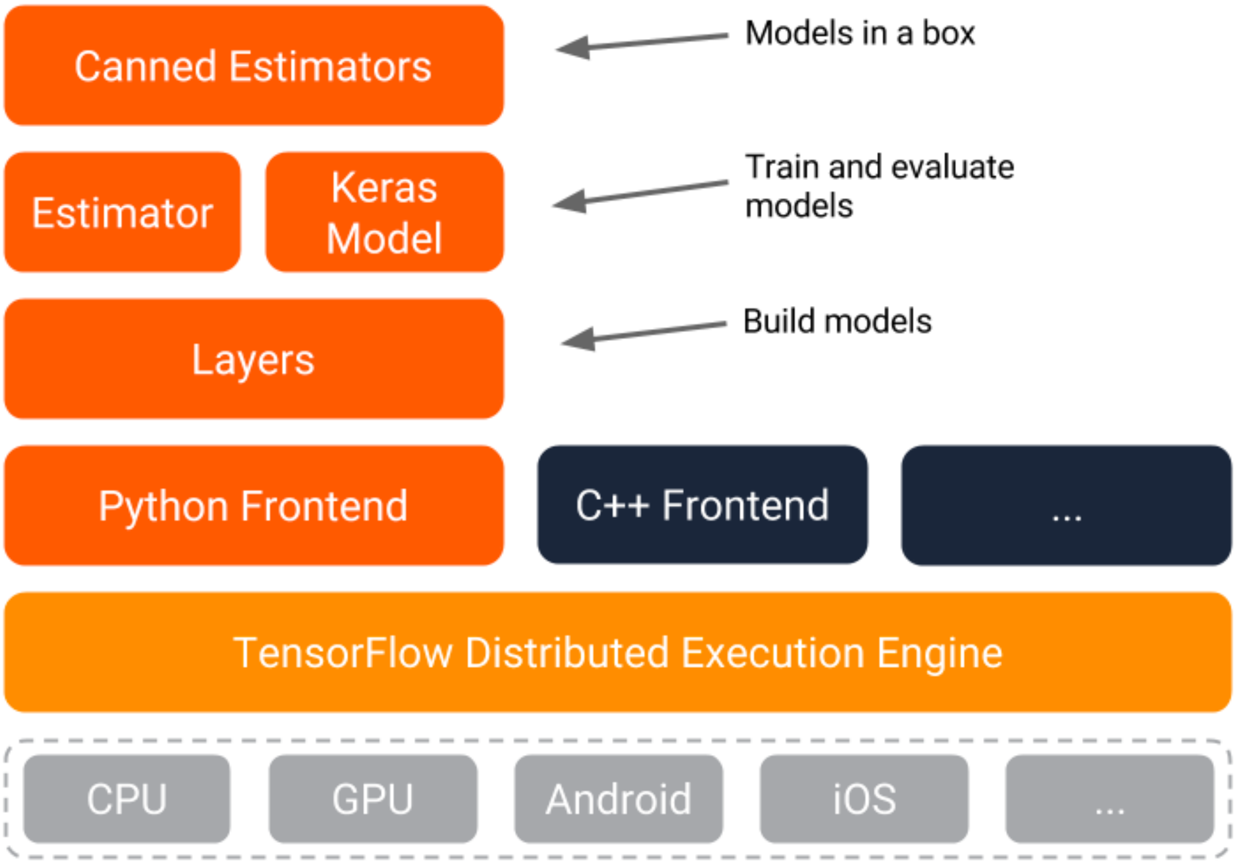
One small piece of what you'd have to do without it…

333 lines (287 sloc) | 11.9 KB   Raw Blame History

```python
from itertools import count
from functools import reduce
from .tracer import trace, primitive, toposort, Node, Box, isbox, getval
from .util import func, subval

# -------------------- reverse mode --------------------

def make_vjp(fun, x):
    start_node = VJPNode.new_root()
    end_value, end_node =  trace(start_node, fun, x)
    if end_node is None:
        def vjp(g): return vspace(x).zeros()
    else:
        def vjp(g): return backward_pass(g, end_node)
    return vjp, end_value

def backward_pass(g, end_node):
    outgrads = {end_node : (g, False)}
    for node in toposort(end_node):
        outgrad = outgrads.pop(node)
        ingrads = node.vjp(outgrad[0])
        for parent, ingrad in zip(node.parents, ingrads):
            outgrads[parent] = add_outgrads(outgrads.get(parent), ingrad)
    return outgrad[0]

class VJPNode(Node):
    __slots__ = ['parents', 'vjp']
    def __init__(self, value, fun, args, kwargs, parent_argnums, parents):
        self.parents = parents
        try:
            vjpmaker = primitive_vjps[fun]
        except KeyError:
            fun_name = getattr(fun, '__name__', fun)
            raise NotImplementedError("VJP of {} wrt argnums {} not defined"
                                      .format(fun_name, parent_argnums))
        self.vjp = vjpmaker(parent_argnums, value, args, kwargs)

    def initialize_root(self):
        self.parents = []
        self.vjp = lambda g: ()

primitive_vjps = {}
def defvjp_argnums(fun, vjpmaker):
    primitive_vjps[fun] = vjpmaker

def defvjp_argnum(fun, vjpmaker):
    def vjp_argnums(argnums, *args):
        vjps = [vjpmaker(argnum, *args) for argnum in argnums]
        return lambda g: (vjp(g) for vjp in vjps)
    defvjp_argnums(fun, vjp_argnums)

def defvjp(fun, *vjpmakers, **kwargs):
    argnums = kwargs.get('argnums', count())
    vjps_dict = {argnum : translate_vjp(vjpmaker, fun, argnum)
                 for argnum, vjpmaker in zip(argnums, vjpmakers)}
    def vjp_argnums(argnums, ans, args, kwargs):
        L = len(argnums)
        # These first two cases are just optimizations
        if L == 1:
            argnum = argnums[0]
            try:
                vjpfun = vjps_dict[argnum]
            except KeyError:
                raise NotImplementedError(
                    "VJP of {} wrt argnum 0 not defined".format(fun.__name__))
            vjp = vjpfun(ans, *args, **kwargs)
            return lambda g: (vjp(g),)
        elif L == 2:
            argnum_0, argnum_1 = argnums
            try:
                vjp_0_fun = vjps_dict[argnum_0]
                vjp_1_fun = vjps_dict[argnum_1]
            except KeyError:
                raise NotImplementedError(
                    "VJP of {} wrt argnums 0, 1 not defined".format(fun.__name__))
            vjp_0 = vjp_0_fun(ans, *args, **kwargs)
            vjp_1 = vjp_1_fun(ans, *args, **kwargs)
            return lambda g: (vjp_0(g), vjp_1(g))
        else:
            vjps = [vjps_dict[argnum](ans, *args, **kwargs) for argnum in argnums]
```

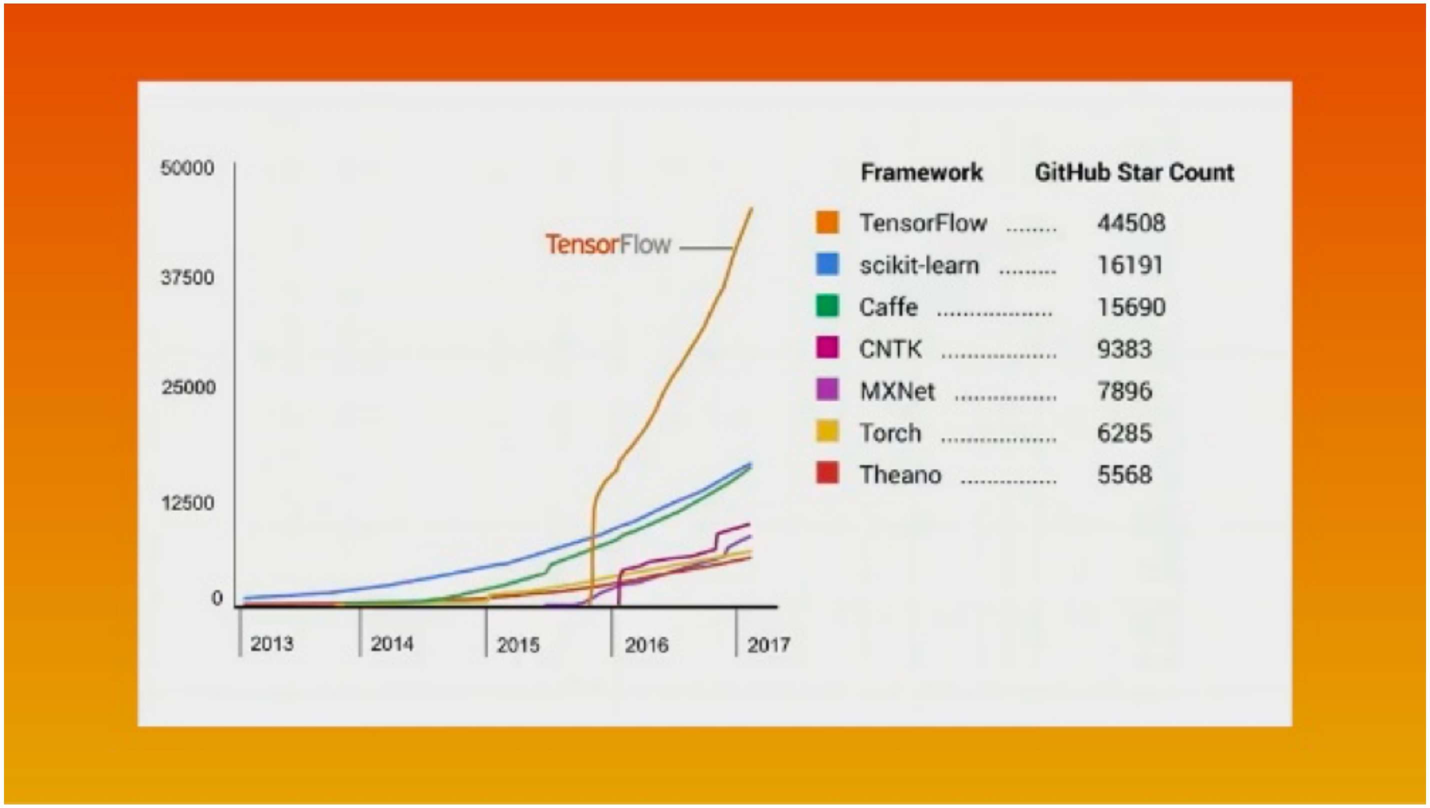
back

TensorFlow makes building deep learning applications easy.



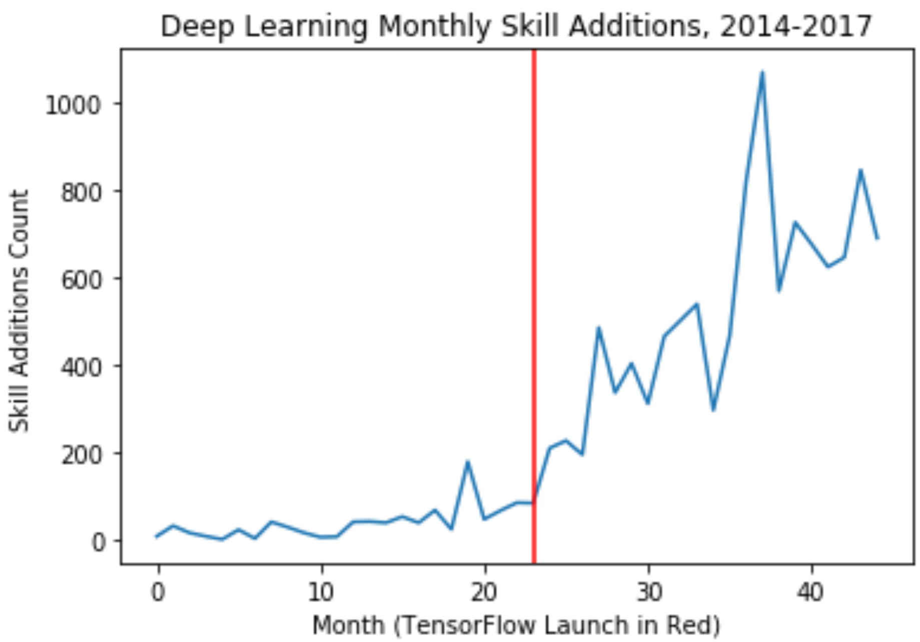
Source: Google

Now the leading deep learning library



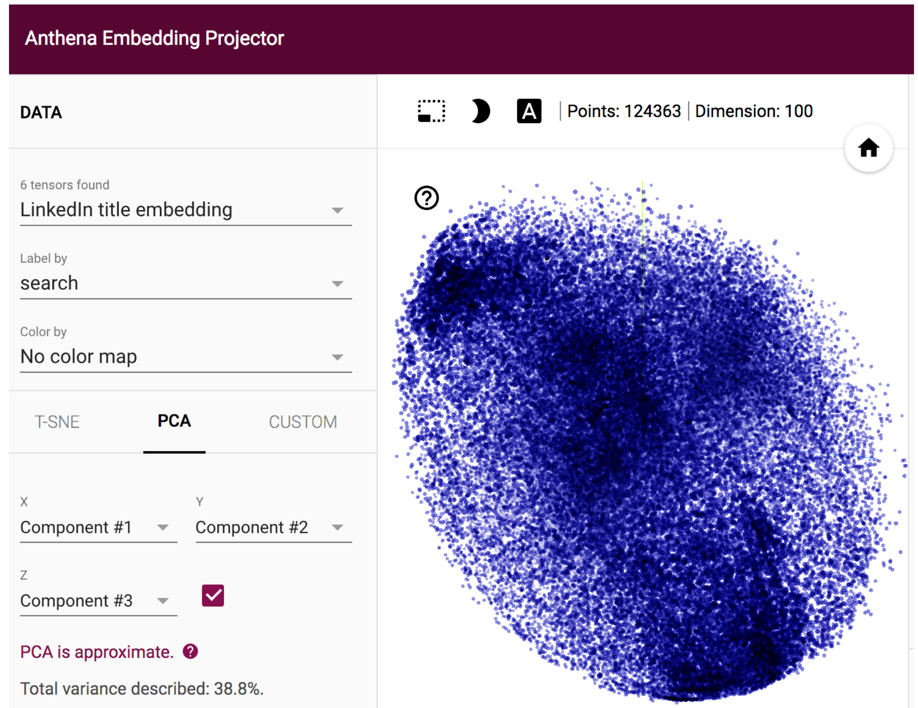
Source: Jeff Dean (Talk at MIT 2018)

LinkedIn Skills data can reveal the labor effects of TF



How do we aggregate the 35,000 skills on LinkedIn?

- Use Machine Learning! We helped to build an *embedding space* of skills into larger interpretable clusters



Some groupings we get

- AI Skills: Machine Learning, computer vision, neural networks, speech recognition, NLP, expert systems, genetic algorithms, reinforcement learning, tensorflow, keras, pytorch
- Data Science: analytics, forecasting, experimental design, probability, tableau, decision trees, R
- Economics: valuation, econometrics, STATA, industrial organization, financial data
- Robotics: motion control, mechatronics, actuators, industrial robots, robocad, robotic design

Where is AI/ML Talent?

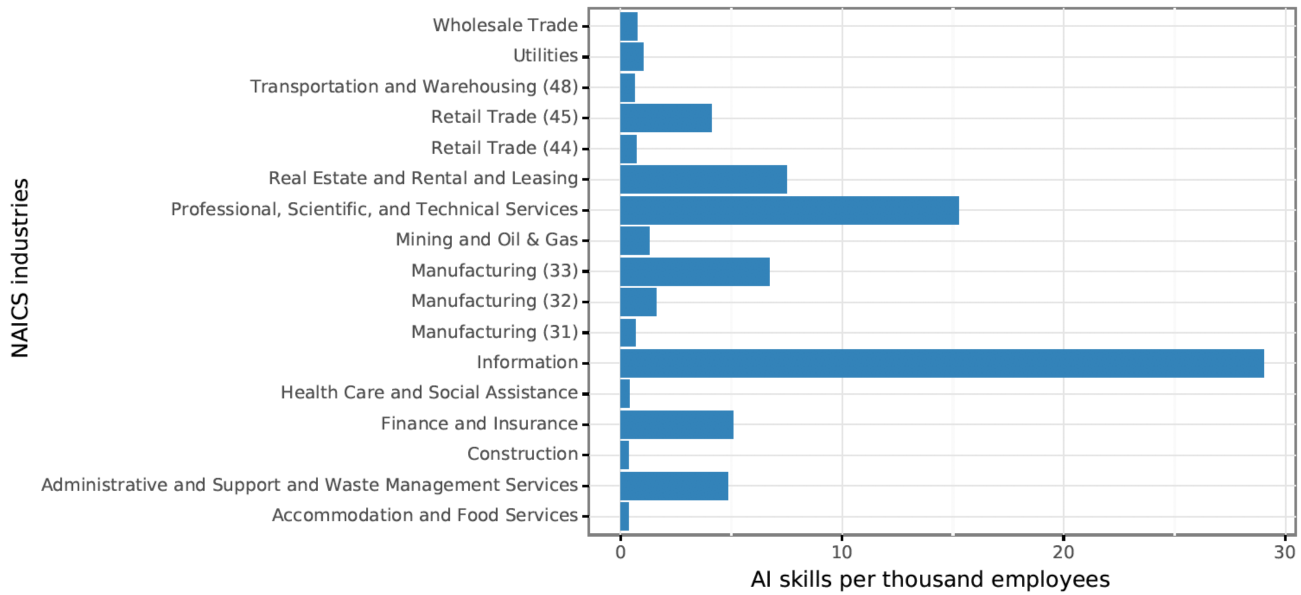


Figure notes: This chart illustrates average AI skills per 1000 employees for a balanced panel of publicly traded US firms in 2017. Industries are categorized as 2-Digit NAICS codes. The Information (NAICS 51) and Professional, Scientific, and Technical Services (NAICS 54) industries have the highest concentration of AI skills. Industries with fewer than 10 firms are omitted.

Source: Tambe, Hitt, Rock, and Brynjolfsson (2019)

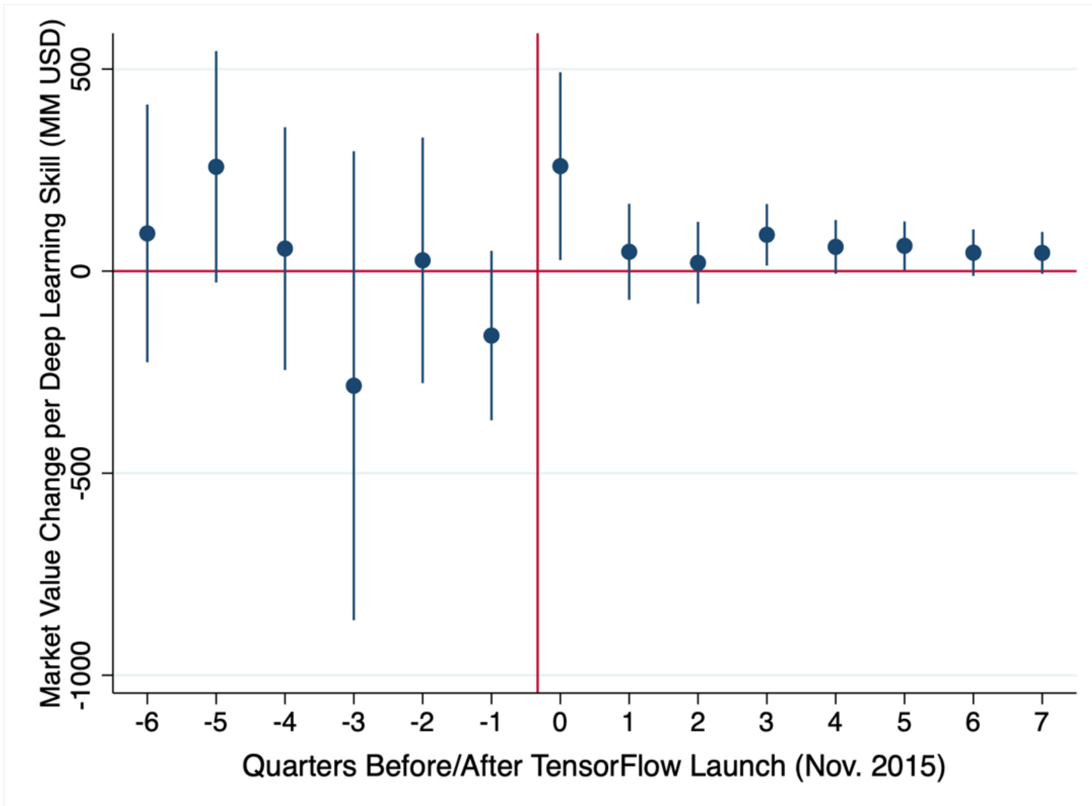
The Prediction Question: Is AI Intangible Capital Priced?

	(1)	(2)	(3)	(4)	(5)	(6)
	Excess MV	Excess MV	Revenues	Revenues	VA	VA
Total Assets			0.08***	0.04***	0.04***	0.03***
			(0.02)	(0.01)	(0.01)	(0.00)
Log(Education)	1362.14*	3455.83***	950.34***	3320.86***	463.66***	952.39***
	(801.55)	(662.64)	(305.39)	(662.14)	(126.65)	(181.67)
Log(AI Skills)	928.79***	6384.37***	-71.35	3332.38***	63.43	1143.62***
	(248.47)	(1259.79)	(151.80)	(673.20)	(39.13)	(220.62)
Log(Data Science)	767.04*	-1460.77*	90.72	-1220.20*	5.86	-590.91***
	(392.41)	(795.77)	(80.55)	(623.73)	(36.80)	(220.66)
Log(Cloud Computing)	189.70	1075.45	31.52	437.49	-26.48	342.57**
	(235.99)	(730.69)	(179.59)	(546.33)	(37.25)	(163.67)
Log(Data Storage)	-5650.68	-103177.93***	-2939.62	-44438.77*	3478.53	-17440.26**
	(14101.90)	(36148.04)	(9126.53)	(22590.43)	(2523.62)	(7331.85)
Log(Digital Literacy)	126.96	454.46	125.64	205.09	10.75	69.92
	(360.17)	(1624.07)	(157.47)	(651.77)	(46.22)	(292.15)
Log(Management)	186.93	-2112.36	143.56	-595.43	191.75**	-202.06
	(631.81)	(1507.33)	(411.27)	(702.91)	(91.97)	(264.65)
Log(Advertising)	293.85	2966.00**	111.14	571.97	6.48	471.31**
	(223.61)	(1293.71)	(183.94)	(638.34)	(34.44)	(205.74)
Firm and Year FE	x		x		x	
Industry-Year FE		x		x		x
R^2	.9392377	.3404959	.985477	.5394115	.9880268	.6634429
N	5,288	5,076	5,288	5,076	5,288	5,076
Clusters	1,322	149	1,322	149	1,322	149

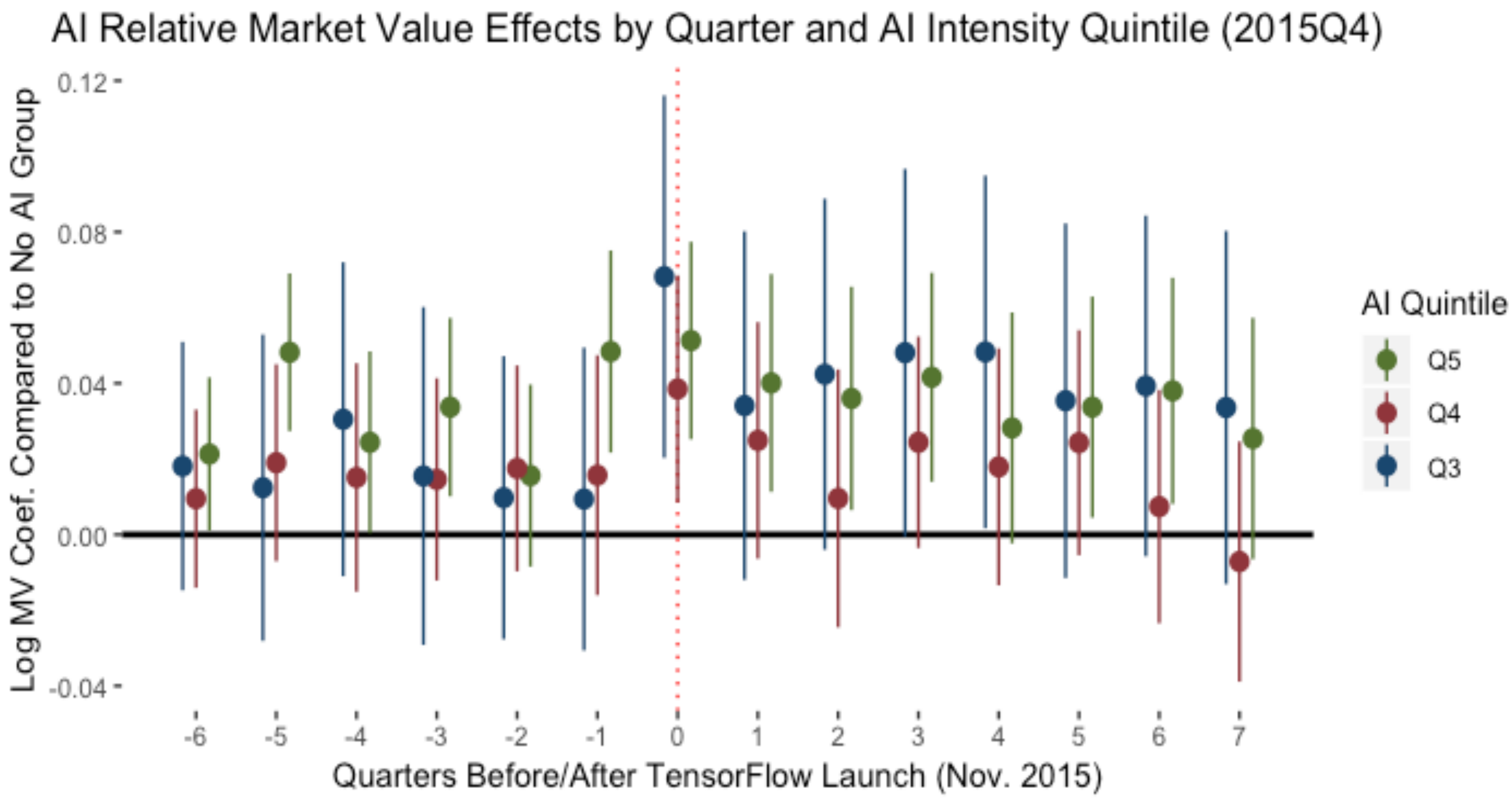
AI-intensive companies are repriced in 2016

Market Value (MM USD)	(1) AI Cluster	(2) +Data Science	(3) +Cloud Computing	(4) +Data Storage	(5) +Digital Literacy	(6) +Bus.Mgmt and Advertising
Log(Total Assets)	7,790***	7,798***	7,798***	7,790***	7,837***	7,832***
	(1,276)	(1,277)	(1,277)	(1,275)	(1,278)	(1,275)
Log(Edu. Years)	-1,582	-1,577	-1,578	-1,562	-1,535	-1,528
	(1,014)	(1,006)	(1,009)	(1,007)	(997.5)	(999.5)
Log(AI Index)	-1,630	-1,616	-1,618	-1,610	-1,601	-1,592
	(1,017)	(1,022)	(1,006)	(1,006)	(1,006)	(1,002)
Log(AI Index x Post TF)	3,299***	3,303***	3,302***	3,305***	3,312***	3,316***
	(712.3)	(712.0)	(715.4)	(715.7)	(716.2)	(719.5)
Log(Data Science Index)		-133.0	-136.8	-104.9	187.4	176.2
		(574.9)	(559.0)	(558.1)	(446.5)	(566.1)
Log(Cloud Computing Index)			16.68	46.40	80.63	81.93
			(397.7)	(398.8)	(389.8)	(419.9)
Log(Data Storage Technology Index)				-29,129	-26,872	-27,442
				(21,191)	(21,370)	(21,357)
Log(Digital Literacy Index)					-639.1	-672.9
					(595.3)	(768.3)
Log(Bus. Management Index)						290.6
						(1,444)
Log(Advertising Index)						-235.1
						(422.4)
Observations	6,440	6,440	6,440	6,440	6,440	6,440
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Deep Learning per Skill Valuation Changes Over Time



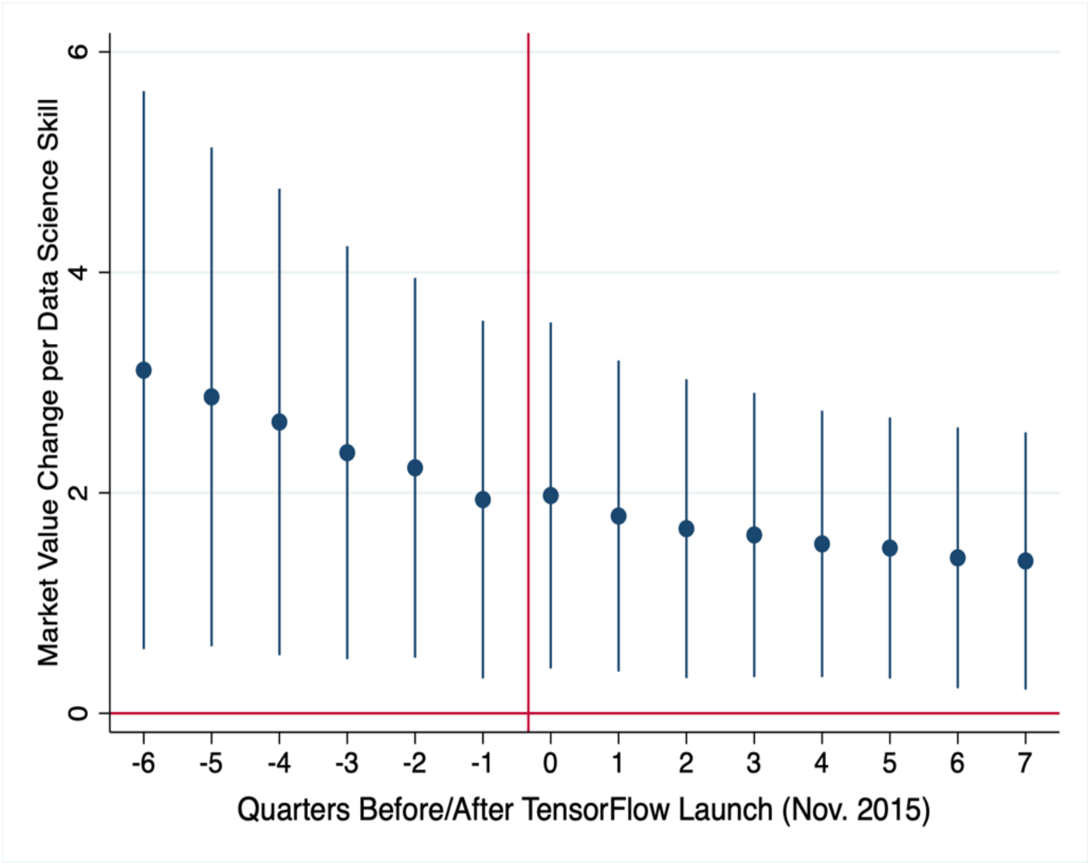
Is the change contemporaneous with the TensorFlow Launch?



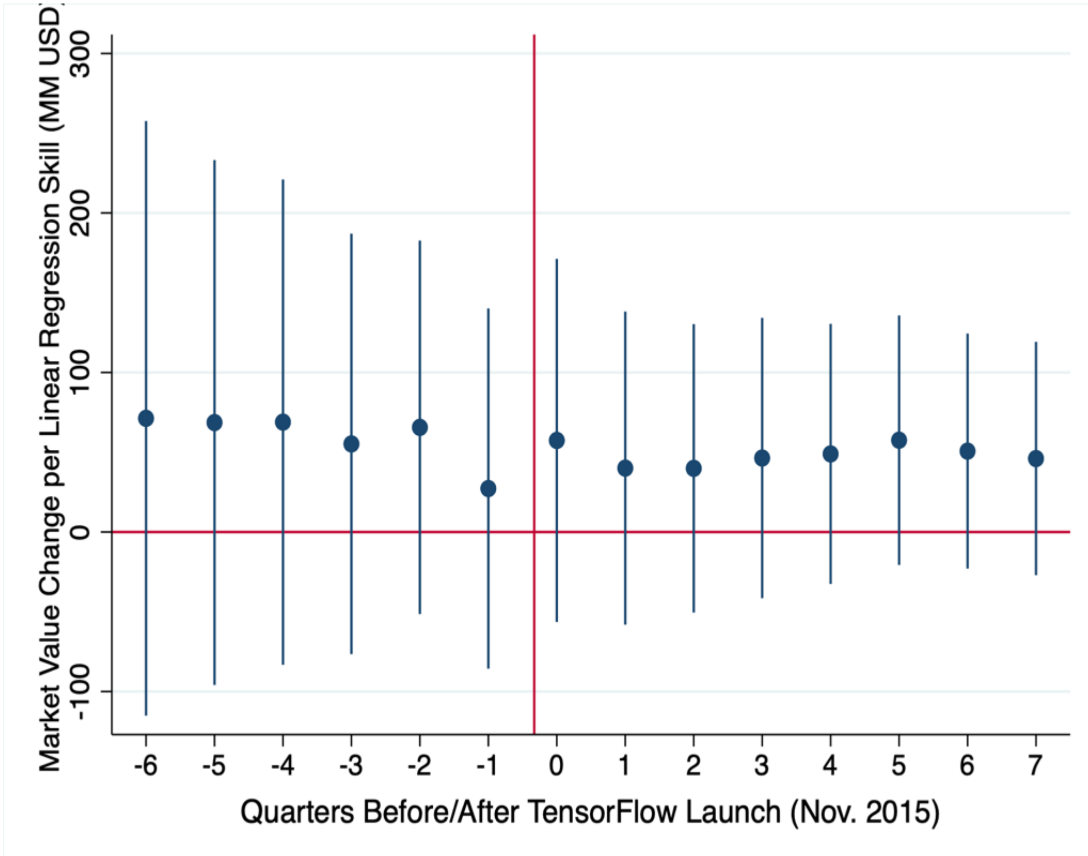
Removing the top quintile of AI-Using firms

Market Value (MM USD)	(1) AI Cluster	(2) +Data Science	(3) +Cloud Computing	(4) +Data Storage	(5) +Digital Literacy	(6) +Bus.Mgmt and Advertising
Log(Total Assets)	5,366***	5,356***	5,360***	5,353***	5,354***	5,370***
	(901.2)	(900.5)	(901.8)	(900.6)	(900.4)	(902.9)
Log(Edu. Years)	-740.1	-745.7	-763.0	-751.7	-751.5	-769.0
	(555.1)	(549.0)	(546.8)	(544.2)	(540.6)	(533.9)
Log(AI Index)	656.7	640.8	608.3	614.3	614.3	603.1
	(643.3)	(649.4)	(647.9)	(649.0)	(648.4)	(650.8)
Log(AI Index x Post TF)	980.0***	974.3***	958.5***	958.4***	958.5***	944.9***
	(290.8)	(288.5)	(289.6)	(289.5)	(289.7)	(287.8)
Log(Data Science Index)		167.3	93.18	116.1	117.8	207.6
		(465.3)	(442.7)	(440.6)	(307.0)	(473.0)
Log(Cloud Computing Index)			332.2	354.6	354.8	378.7
			(284.7)	(285.9)	(275.4)	(309.4)
Log(Data Storage Technology Index)				-21,531	-21,518	-20,920
				(13,919)	(14,125)	(14,272)
Log(Digital Literacy Index)					-3.713	127.4
					(470.2)	(693.5)
Log(Bus. Management Index)						-799.1
						(1,311)
Log(Advertising Index)						352.5
						(356.3)
Observations	5,864	5,864	5,864	5,864	5,864	5,864
Firm and Year FE	Yes	Yes	Yes	Yes	Yes	Yes

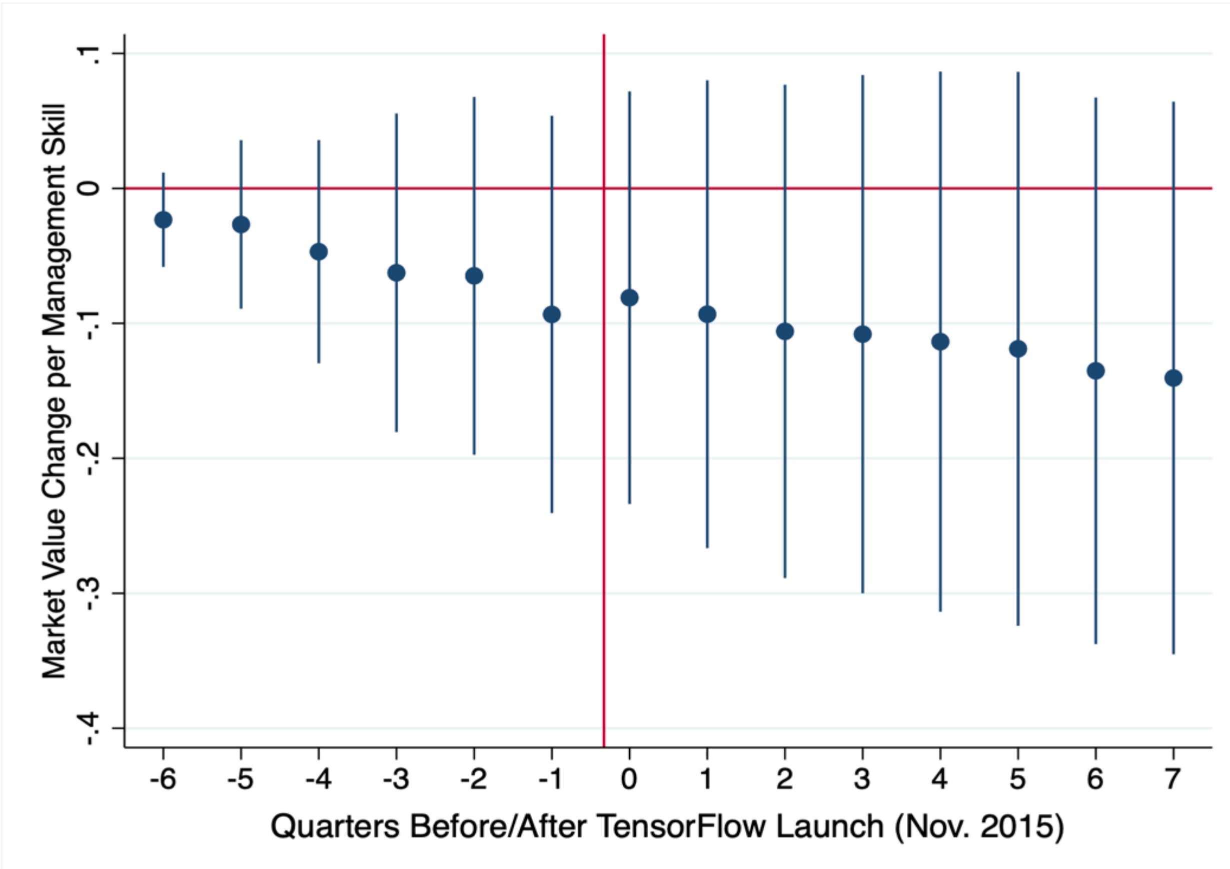
Data Science per Skill Valuation Changes Over Time



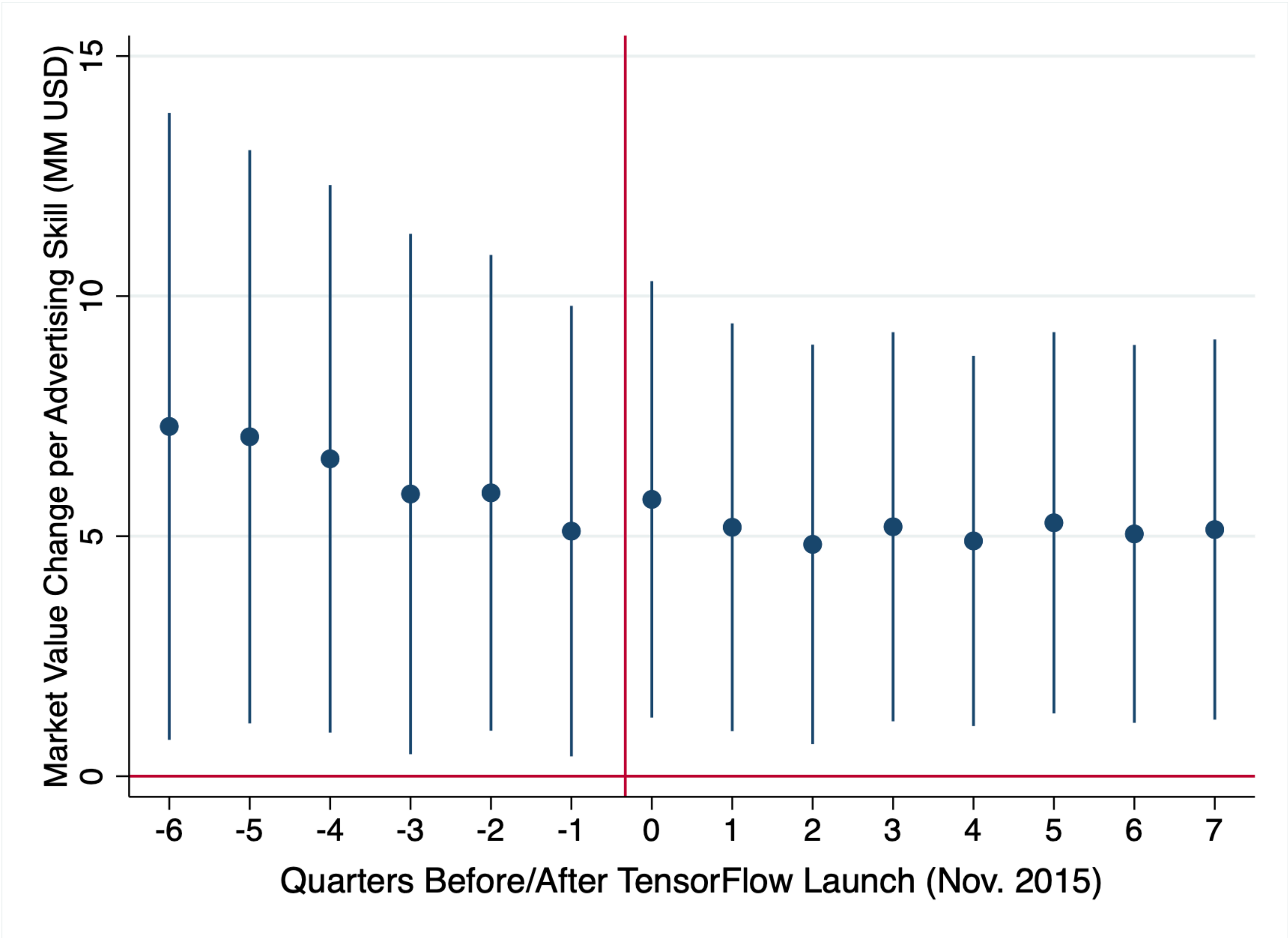
Linear Regression per Skill Valuation Changes Over Time



Management per Skill Valuation Changes Over Time



Advertising per Skill Valuation Changes Over Time



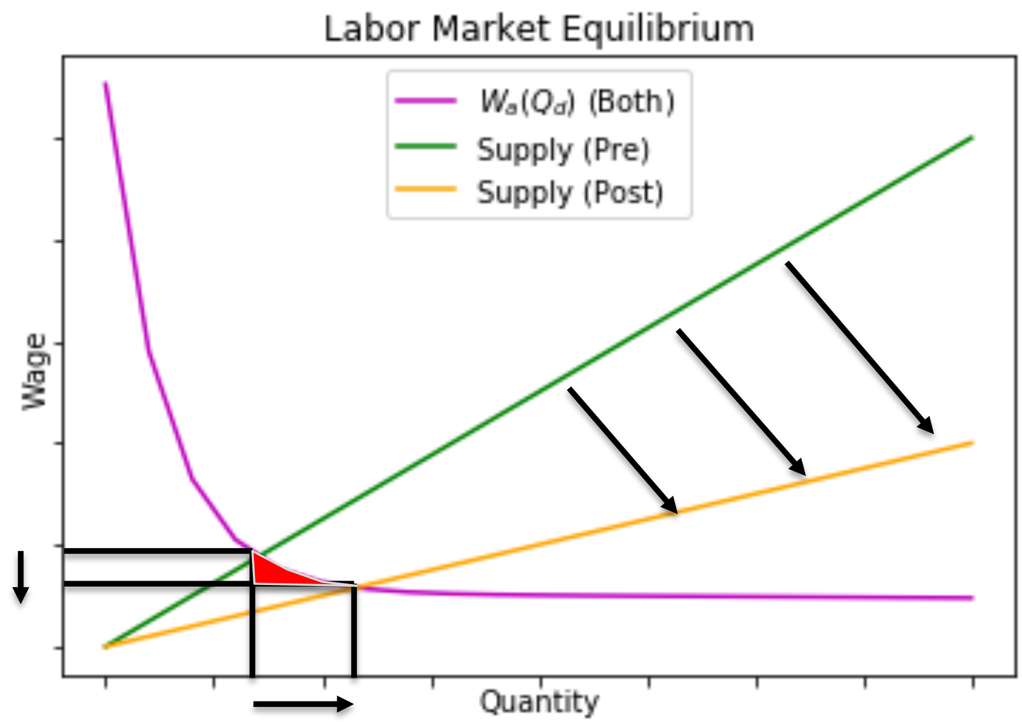
Synthetic Difference-in-Differences Coefficients are similar*

Synthetic Difference-in-Differences Results

	Skill Group	Estimate (Billions USD)	Standard Error (Billions USD)	N Treated
1	Artificial Intelligence	1.6	0.75	605
2	Data Science	0.72	0.41	1262
3	Deep Learning	14.3	4.9	60
4	Cloud Computing	0.93	0.53	943
5	Business Management	0.25	0.39	1357
6	Linear Regression	3.06	1.74	199
7	SML (Median)	0.16	0.3	1037

*Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager. *Synthetic difference in differences*. No. w25532. National Bureau of Economic Research, 2019.

TensorFlow: A technological shock to the expected future talent supply



Who will earn the returns to AI Talent?

- Engineering talent needed to implement new technology
- …but extensive firm co-investment requires to realize its returns
- Punchline: the engineering talent value goes to the company too
    - When the margins change: **4-7% MV increase**
        - Middle firms benefit! TensorFlow was democratizing!
        - Mostly an AI intangible price shock for AI-using firms
- Technological shocks help understand employer-employee relationships

Thanks!

Contact: @danielrock, rockdi@wharton.upenn.edu