Measuring Moore's Law:
Evidence from Price, Cost, and Quality Indexes

Kenneth Flamm
University of Texas at Austin

November 2017

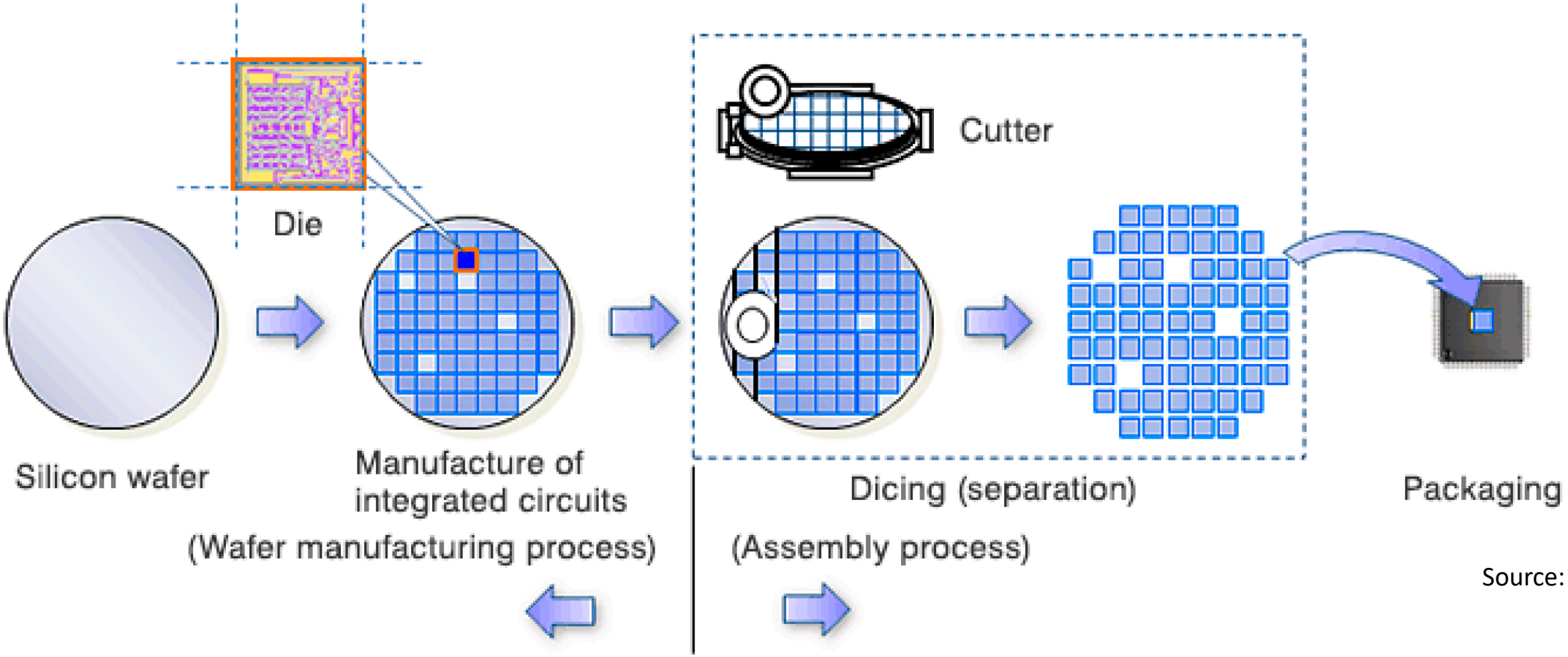
Semiconductors are important

- See Jorgenson, Oliner, Sichel, Aizcorbe, Byrne, Corrado, Doms etc.
- Technological foundation of IT industries
- Historically, measured poorly in official price indexes
- Measurement situation got better for a while in late 1990s, early 2000s
- Problem is now getting worse, not better
- Private data is getting worse
  - Probably related to **industry consolidation (more below)**
- US public data non-existent

“Moore’s Law” is Industry Bumper Sticker for One Very Important Source of Technological Innovation in Semiconductor Manufacturing

- Refers to effect on manufacturing cost of regular, predictable adoption of new fabrication technology which shrank chip size as patterned on surface of silicon wafer
- Predictably lowered cost per transistor by 20-30% annually for commodity chips
- Lowered cost for most types of chips
  - Most chip designs eventually migrated to new technology eventually
- Not the only source of innovation in semiconductor manufacturing
- Important innovation in semiconductor circuit design and functionality
- Will discuss evidence suggesting Moore’s Law has recently slowed, and is effectively approaching end
- Will link to available data for industry

Semiconductor Manufacturing



Source: Hitachi

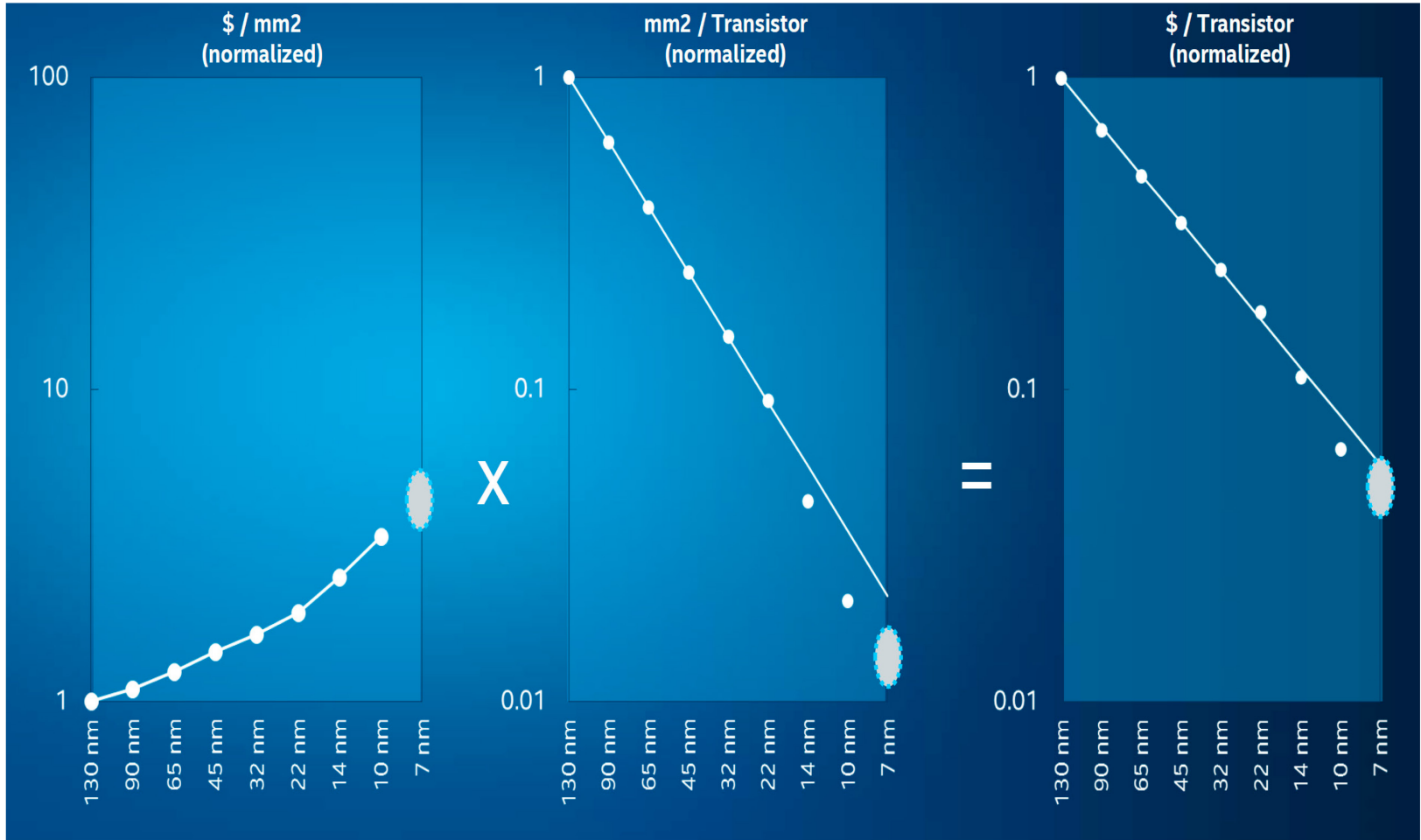
Fabrication

Assembly, packaging, final test

“Moore’s Law” is Shorthand for Economics of Semiconductor Industry

- mfg cost: $/element = $\frac{\text{wafer processing cost}}{\text{area yielded silicon}} \times \frac{\text{silicon area}}{\text{element}}$

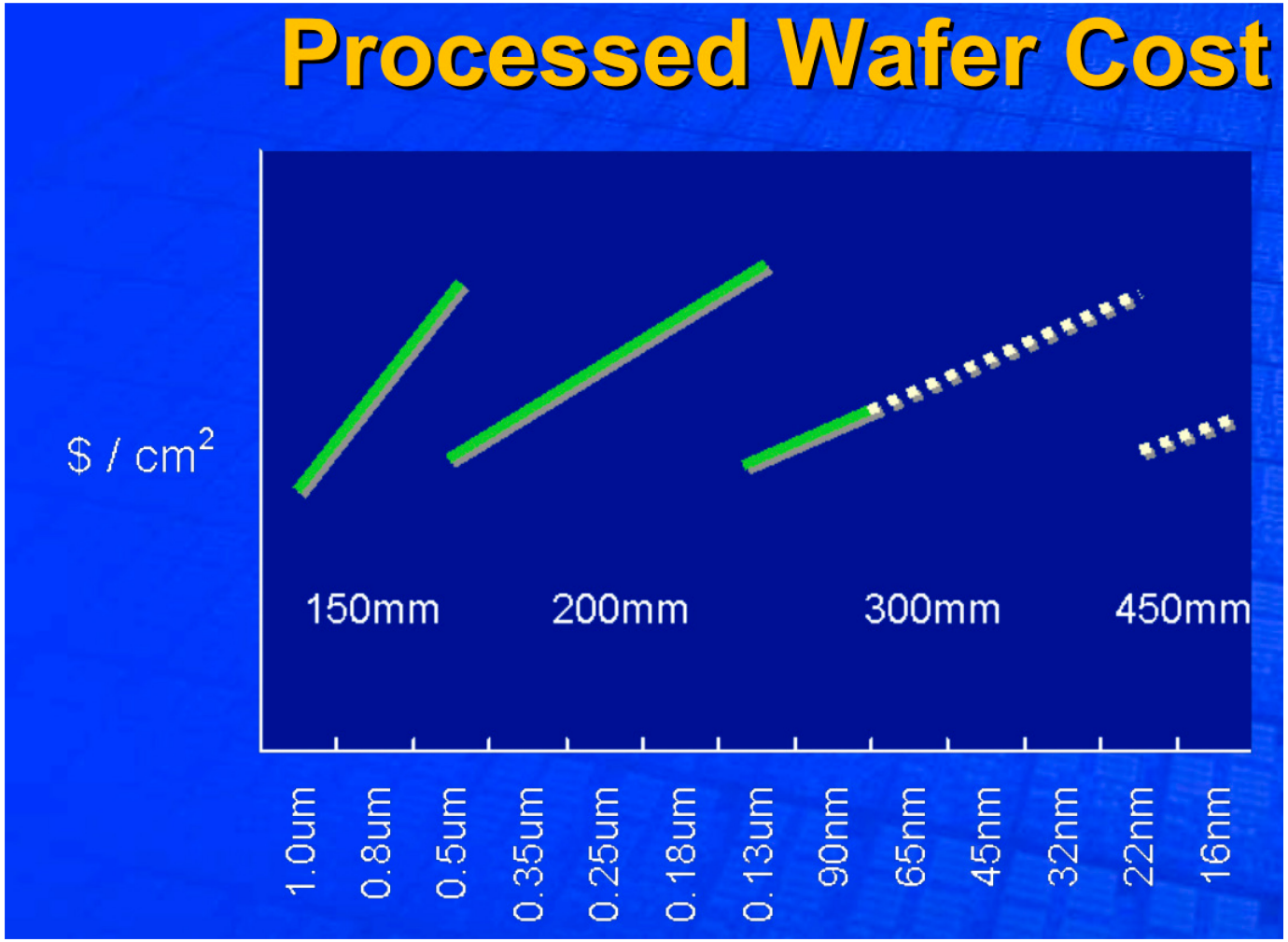
- Intel’s 2015 version:



Wafer Processing Cost

- Chips are fabricated on silicon wafers
- Processing cost / area ≈ close to constant, in historical long run
  - Index of patterning process is "critical feature size"
    - Measured in nanometers (nm)
- But 450mm wafers didn't happen
- Processing cost/ area increasing
  - Since 90nm technology "node"

Intel view, 2005:



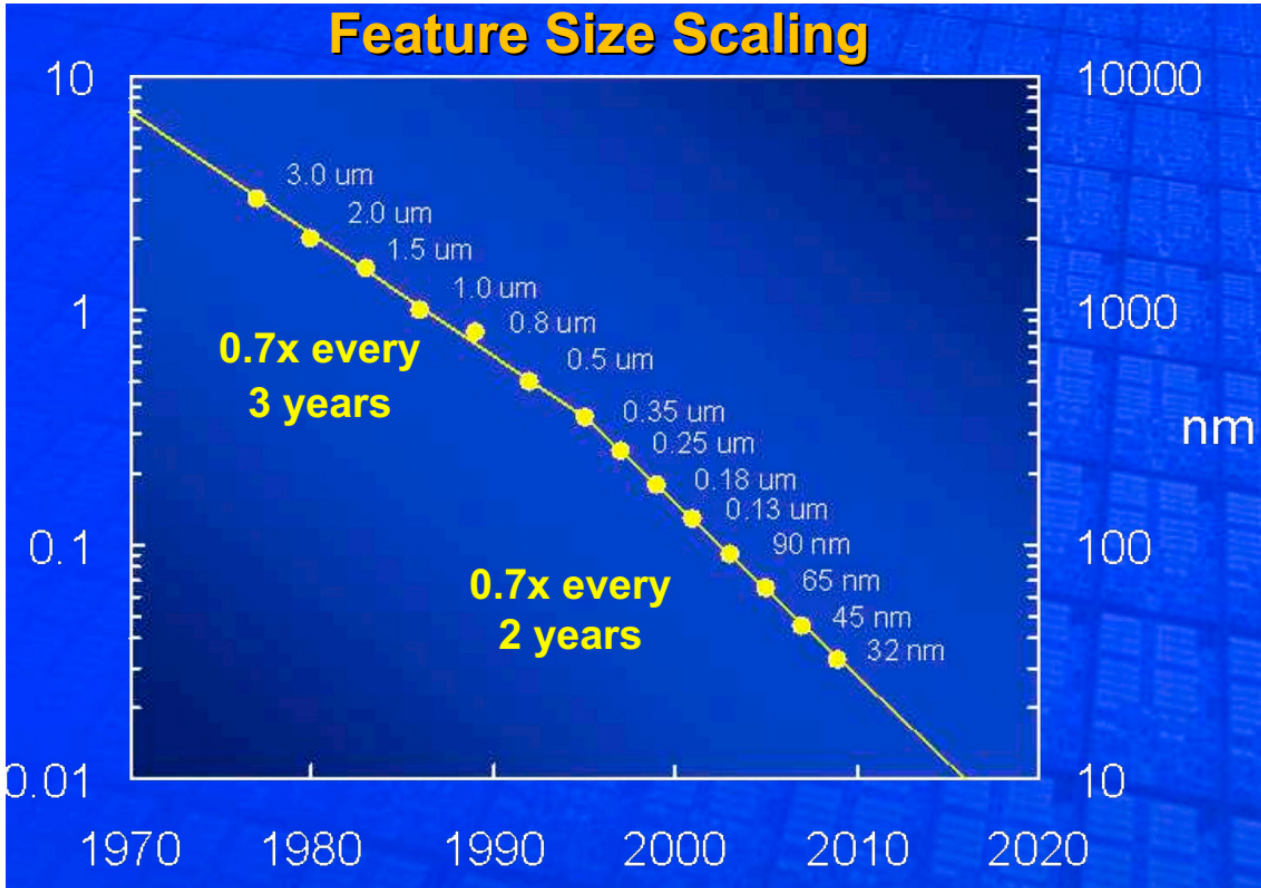
Silicon area/transistor

- Silicon wafer area/transistor ≈ 50% reduction every new **process technology node**

- New process (technology node) every every N years
  - N ≈ 3 prior to 1995
  - **N≈ 2 1995-2014**

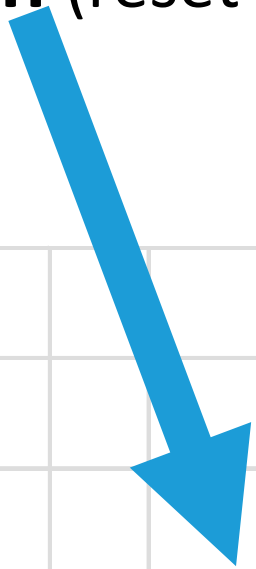
- 2-year cadence didn't last forever!
  - **N ≈ 4, 2014+**

- Intel slide, 2005➔



Implications: the Moore's Law calculator

- Assume **historical pattern** (reset after new wafer size, cost ~ constant)



					New	Annual Cost	
					tech node		
					every:	Change	
% change in cost transistor			-50%		2	-29%	Post '95
	w/ new tech node				3	-21%	Pre '95
					4	-16%	
					years		

But what **if**?

…last wafer size “reset” occurred at 90 nm-130 nm transition ~ 2001-2004

					New tech node every:	Annual Cost Change
% change in cost transistor:			-30%		2	-16%
	w/ new tech node				3	-11%
					4	-9%
					years	

And what if?

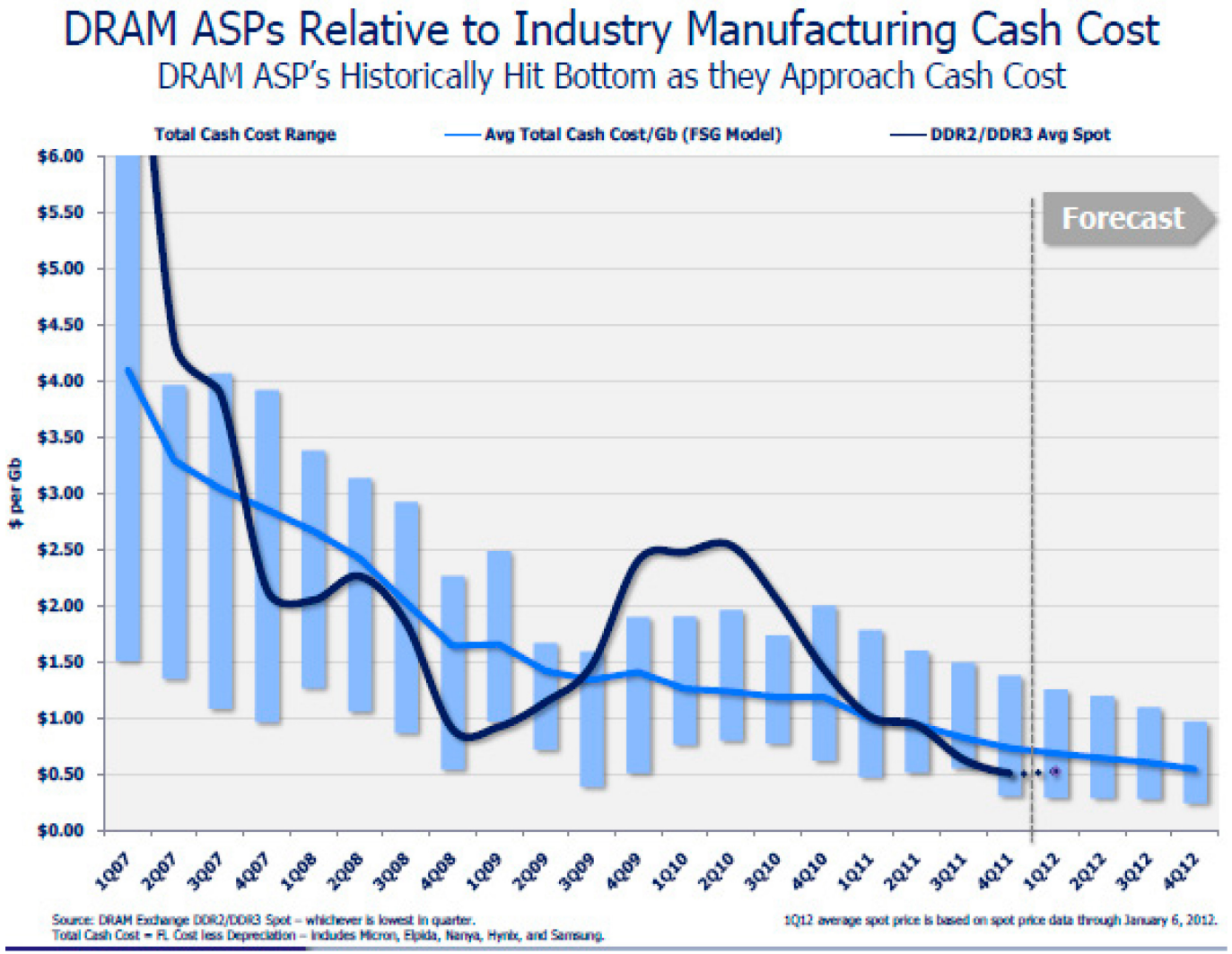
					New	Annual Cost
					tech node	
					every:	Change
% change in cost transistor:			-20%		2	-11%
	w/ new tech node				3	-7%
					4	-5%
					years	

We may be here!

Implications for Price Measurements

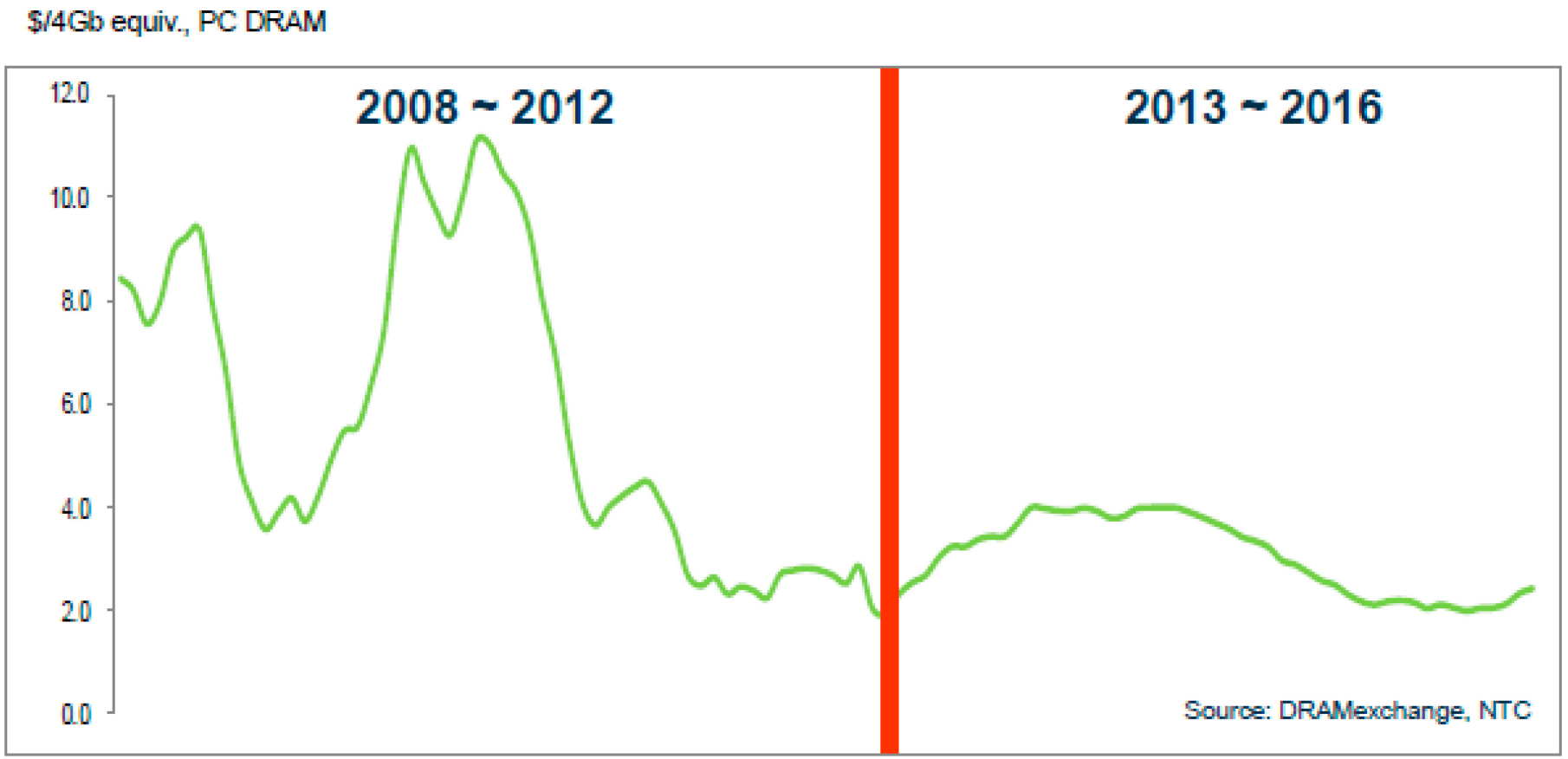
- One very simple way to think about it
- ΔP = Pass-through rate x ΔCost
  - Pass-through = 1 in perfectly competitive industry, constant returns
  - < = > 1 in imperfectly competitive industry
  - Frequently measure empirical pass-through rates around 1 in electronics industries
- Benchmark ideal for "pretty competitive" industry:
  - Price decreases reflect cost declines
  - Pass-through close to 1 would imply 20-30% decline in price of given electronic circuit design if moved to to leading edge technology generation
  - Additional innovation would imply even larger declines in quality-adjusted prices
  - ➔Moore's Law as "floor" on average quality-adjusted price decline: **20-30%**

Price Eventually Approaches Cost With Intense Competition: DRAM Industry Case



Source:
Micron Technology

DRAM Pricing Fluctuation Stabilized



- Limited DRAM supply expected in 2017, leads to healthy market condition

Source: Nanya Technology

But DRAM competition now also less intense!

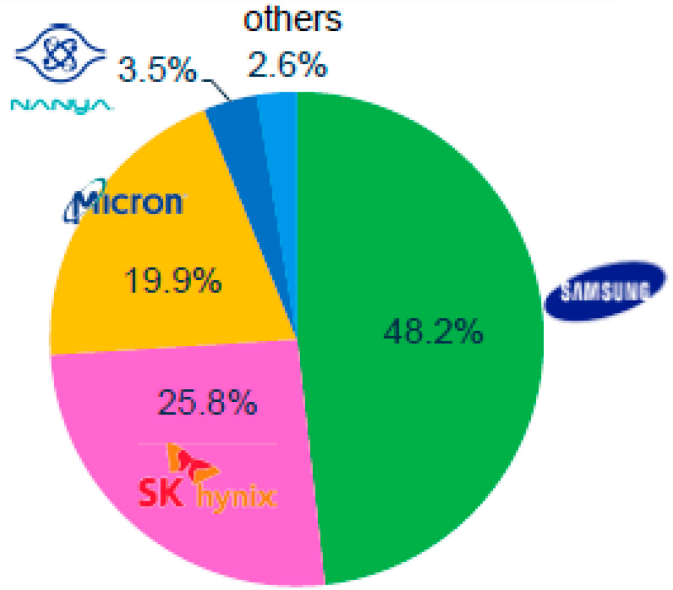
Oligopoly in DRAM Market

- From 2013, DRAM market fluctuation stabilized

1990 (more than 16 suppliers)

Consolidation

2017 (3 plus major suppliers)



Source: Gartner, NTC

Source: Nanya Technology

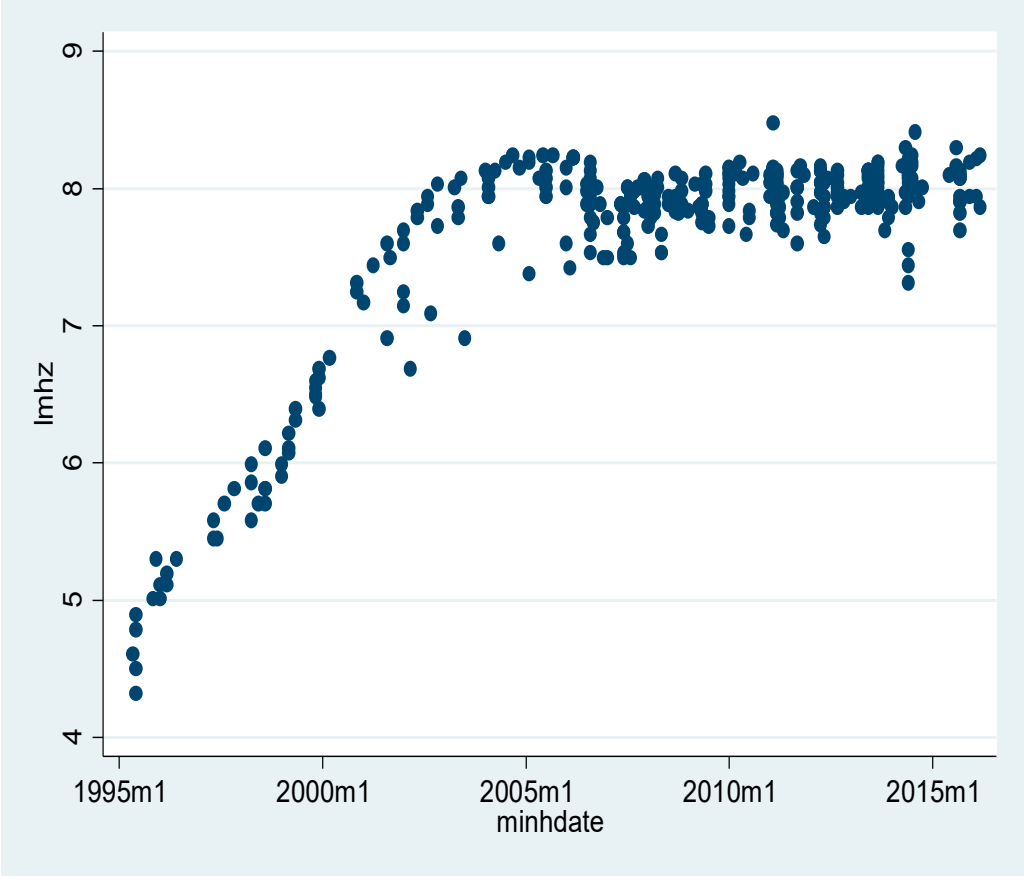
But wait, there's more…

- Pure reduction in manufacturing cost shouldn't necessitate any special quality adjustment in measuring price index, but…
- Smaller transistors also switched ***faster***
- Smaller transistors also ***drew less power***
- Cheaper transistors and ***smaller size*** made it economic to add ***additional functionality*** to chip
- **Additional benefits add value, increase decline in quality-adjusted price, make quality-adjustment necessary**

And…

- **Faster** ended in 2004
- ***Lower power*** now requires tradeoffs
  - Don’t get it for free with smaller/denser any more
- ***Smaller*** still going on but at much slower rate
  - 4 years to next tech node!
- What about price and cost?

Log(Processor Speed)



Evidence from price indexes

Do we see a deceleration in chip prices?

Memory chips

- Micron Tech:

--"technology-driven growth slows due to scaling limits"

--16nm will be last technology node for flash memory

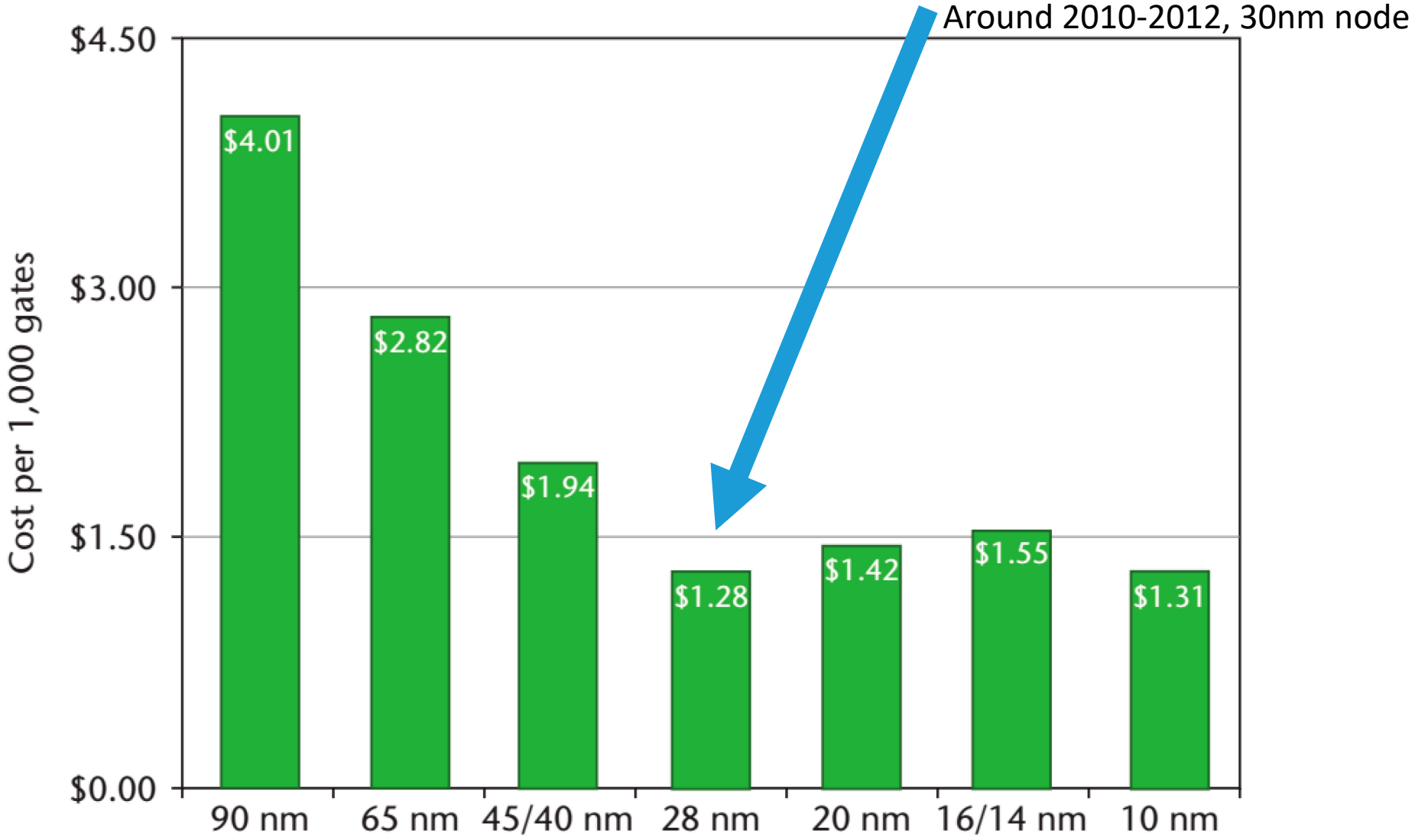
		Compound Annual Decline Rate					
		Flamm-Aizcorbe DRAM Composite	BoK $EPI DRAM	BoK $EPI Flash	BoK DRAM PPI	BoK Flash PPI	BoJ Chain-Wtd MOS Mem PPI
1974:1-1980:1		-45.51					
1980:1-1985:1		-43.45					
1985:1-1990:1		-24.74					
1990:1-1995:1		-17.40	-10.81				
1995:1-1999:4		-46.37	-44.28		-33.26		
1999:4-2005:1			-28.94	-31.28	-31.76		-24.04
2005:1-2011:4			-37.94	-26.92	-30.65	-29.28	-28.79
2011:4-2016:4			2.33	-12.70	-1.42	-5.76	-13.57

Outsourced semiconductor manufacturing (1/4 – 1/3 of industry output)

- Fabless design companies major players now in US chip industry
  - Qualcomm, Broadcom, Nvidia, AMD, etc.
  - Quality-adjusted price index for chips fabricated at “foundries”
  - Byrne, Kovak, & Michaels, 2016

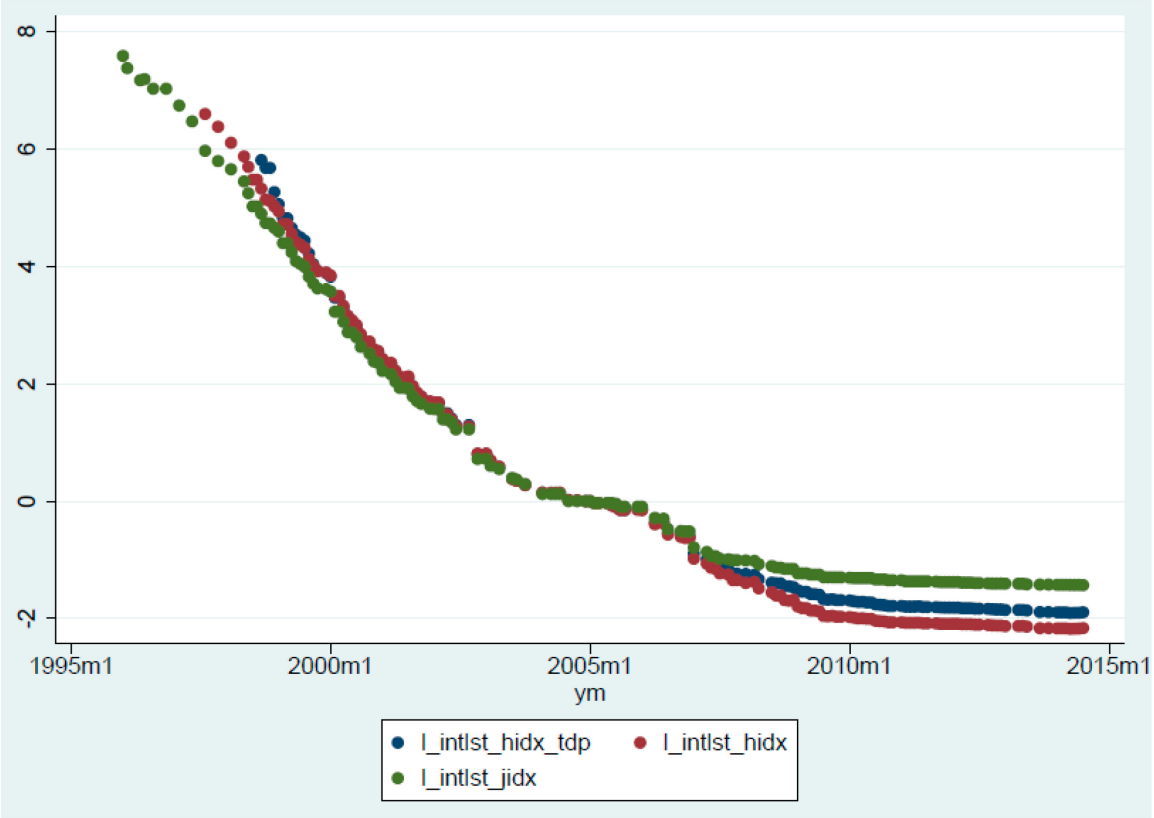
	Annual Index	% Rate of Change
2004	100	
2005	83.89521	-16.1048
2006	74.75891	-10.8901
2007	65.93704	-11.8004
2008	57.89118	-12.2023
2009	52.95437	-8.52774
2010	48.67003	-8.09062

Fabless Chip Designers Now Say Transistor Price is Pretty Flat



Are Intel Processors an Exception?

- Doesn’t look that way in Intel 1K tray list price data sheets
- If you take Intel tray chip prices and do a hedonic price regression on a very complete set of chip characteristics, you get something like:



Similar Pattern in Other Processor Price Indexes

- Note: improved retail and list price indexes coming soon

		Compound Annualized Decline Rate						
		Intel Tray Price				Producer Price		Retail
		Hedonic, no TDP	**Hedonic with TDP**	GeoMean Matched Mocel		Microprocessor PPI		GeoMean Matched Model
1998m9-2001m10		-68.3%	**-73.0%**	-65.0%		-57.5%		
2001m10-2004m2		-50.5%	**-50.1%**	-48.2%		-46.6%		-34.0%
2004m2-2006m1		-14.4%	**-13.8%**	-10.7%		-25.2%		-11.1%
2006m1-2009m1		-42.1%	**-36.9%**	-31.5%		-29.0%		-24.2%
2009m1-2010m11		-13.7%	**-13.6%**	-6.2%		-22.7%		-11.3%
2010m11-2014m7		-2.7%	**-2.9%**	-2.2%		-3.7%		

Counter-point (Byrne, Oliner, Sichel, 2015-6)

- Intel list price data is unreliable after 2006
- Trim post-introductory list price data after 2006, only use initial prices
- Use SPEC CPU Benchmark scores instead of detailed chip characteristics in hedonics
- Get **~40% annual decline throughout post-2000 period**
- **Moore's Law unchanged**?

Could Intel be an Exception?

- Reasons to think they could be:
  - **Scale economies at company level…unlikely**
    - $5-$10B for a leading edge fab now
    - Only 4 companies in world currently investing in leading edge fab technology
      - Intel, Samsung, TSMC, Global Foundries
      - Latter two are “pure” foundries, aggregating outsourced designs of others
    - But we just saw what was happening to foundry prices, so not a good explanation
  - **Scale economies at product level…definitely**
    - Fixed design and photomask costs have increased exponentially at recent tech nodes
    - Fancy “computational lithography” with multiple photomask steps for single chip feature pattern
      - immersion in liquid, phase shift masks, computer modeling of lens system now required at recent tech nodes
    - Masks extremely expensive, twice as many process steps required at 22nm vs. 90nm (per Intel)
    - Intel has enormous volumes– maybe 300-400 million processors a year (in 2014) using a small number of basic designs/mask sets
    - Smaller fry do not have this advantage

Is using SPEC instead of characteristics in quality-adjustment hedonics a good idea?

- No, Sawyer and So (2017) show BOS ~ constant decline rate after 2000 is result of use of SPEC Benchmarks ***instead*** of CPU characteristics in hedonic model, not trimming of incumbent models from sample
  - Reject exclusion of characteristics using statistical tests
  - Get single digit decline rates with hedonic model including characteristics, for recent years
- No, economic theory tells us effects of chip characteristics on BOTH demand and cost lead to inclusion in hedonic equation
  - Even if SPEC perfectly represents chip quality on demand side, SPEC doesn't measure effects of different chip characteristics on cost
- No, benchmark scores are almost perfectly explained by 5 chip characteristics + chip architecture indicator variables
  - As predicted by computer architecture literature
  - SPEC benchmarks look just like fixed-weight combinations of characteristics
    - ***Weights are quite different for different benchmarks***
  - Fixed weights won't be right if market demand shifts across benchmark app types

SPEC benchmarks almost perfectly explained by small set of processor characteristics

Table 11 Log of Median SPEC 2006 Benchmark as Function of Processor Characteristics
Five Characteristics Model

Dependent variable is log of median computer system score for particular processor model

	SPECf06	SPECi06	SPECfr06	SPECir06
lproc	0.265***	0.150***	0.497***	0.439***
	(0.0351)	(0.0376)	(0.0840)	(0.0672)
lcache	0.0788**	0.0582**	0.164**	0.137***
	(0.0254)	(0.0191)	(0.0591)	(0.0295)
lcores	0.143***	0.0446	0.559***	0.678***
	(0.0258)	(0.0263)	(0.0527)	(0.0297)
lvcore	0.0603***	0.0315***	0.0963***	0.149***
	(0.0152)	(0.00451)	(0.0152)	(0.00787)
lmaxmhz	0.453***	0.692***	0.0151	0.334***
	(0.0652)	(0.0551)	(0.114)	(0.0644)
Microarchitecture dummies	Y	Y	Y	Y
Observations	331	340	449	454
R-squared	0.988	0.985	0.990	0.994
N_clust	30	30	28	28

Cluster robust standard errors in parentheses, clustered on Intel microarchitecture.
* p<0.05, ** p<0.01, *** p<0.001

Implications for choice of characteristics in processor hedonic price equation

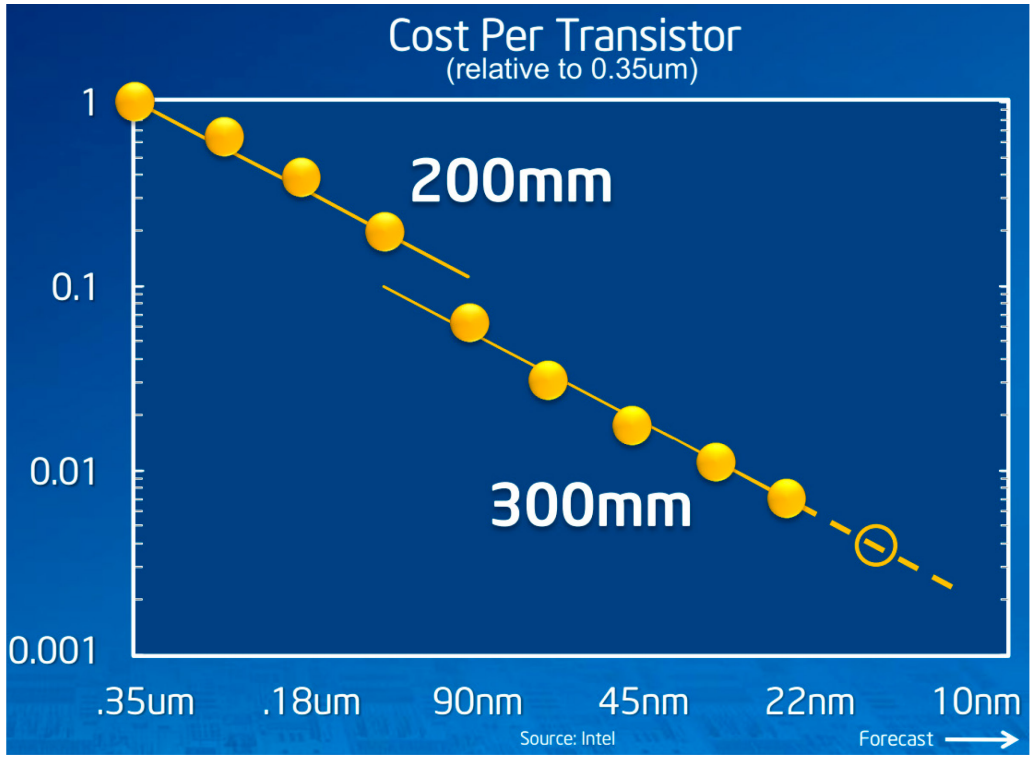
- Characteristics more flexible, accomodate demand shifts across user application types
- Use microarchitecture dummy variables, also capture fabrication cost differences
  - Each Intel microarchitecture produced using a single fabrication technology node
- Good choices for characteristics for hedonic price adjustment
  - 5 chip characteristics + chip architecture indicator variables
  - Power draw + virtualization hardware capability + graphics capability
  - Additional 3 characteristics have no direct effect on SPEC benchmarks
  - But important on demand side to specific groups of users
  - Additional capabilities—virtualization, low power draw, graphics—also affect processor cost
- Different coefficients on different SPEC benchmarks suggest desktop, mobile, server processor groupings are useful disaggregation for price measurement

Evidence From Cost Indexes

Is Intel an Exception?

- 2012 Intel Answer: Maybe not…
- Fabrication cost per transistor per Intel Investor Meeting, 2012:

		Compound Annual Decline Rate		
		Otellini, 2012		
		Wafer Size		
Intro Date	Tech Node	200mm	300mm	
1995q2	350			
1997q3	250	-17.1		
1999q2	180	-22.8		
2001q1	130	-32.3		
2004q1	90		-31.5	
2006q1	65		-30.1	
2007q4	45		-27.9	
2010q1	32		-17.9	
2012q2	22		-18.3	



Micron Technology Info on DRAM Production Costs:

	Cost/Gb
• Pre-2012	-30%/Yr
• 2012-2013	-25-30%/Yr
• 2013-2015	-15-20%/Yr
• 2015-2017	-10-15%/Yr

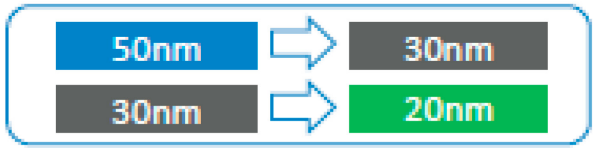
Source: Calculations based on various Micron Technology public investor presentations, 2012-2017.

Why?

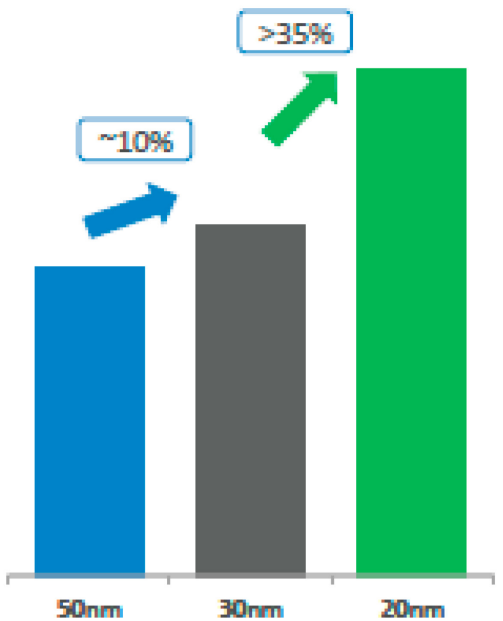
Impacts of DRAM Process Complexity

- Large increase in number of process steps to enable shrink
- Conversion CapEx scales with the number of steps
- Significant reduction in wafer output per existing cleanroom area

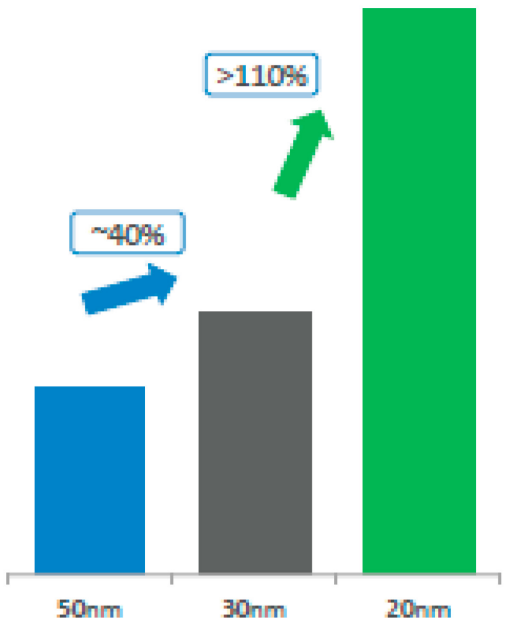
Complexity comparison for enablement of ~100% bits/wafer increase



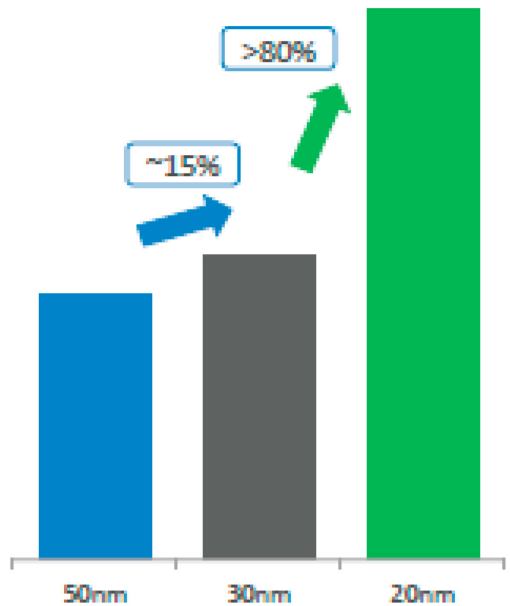
Number of Mask Levels



Number of non-Litho Steps per Critical Mask Level



Cleanroom Space per Wafer Out

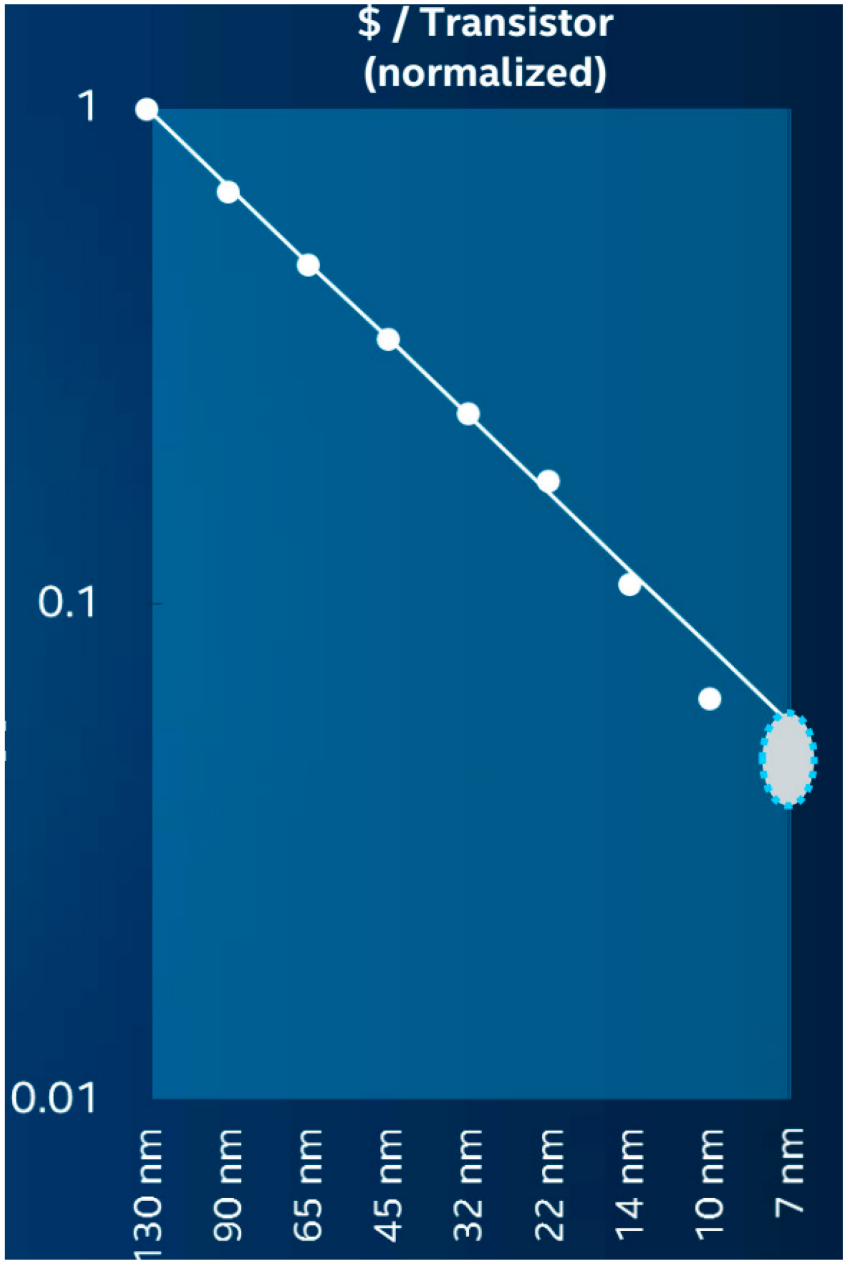


Is Intel an Exception?

- 2015 Intel answer: YES!
  - Switch from empirical to theoretical transistors/mm2 responsible?

		Compound Annual Decline Rate		
		Otellini, 2012		Holt, 2015
		Wafer Size		
Intro Date	Tech Node	200mm	300mm	300mm?
1995q2	350			
1997q3	250	-17.1		
1999q2	180	-22.8		
2001q1	130	-32.3		
2004q1	90		-31.5	-12.0
2006q1	65		-30.1	-15.6
2007q4	45		-27.9	-18.1
2010q1	32		-17.9	-14.2
2012q2	22		-18.3	-13.0
2014q3	14			-19.2
2017q4?	10			-21.1

Intel Investor Meeting, 2015:



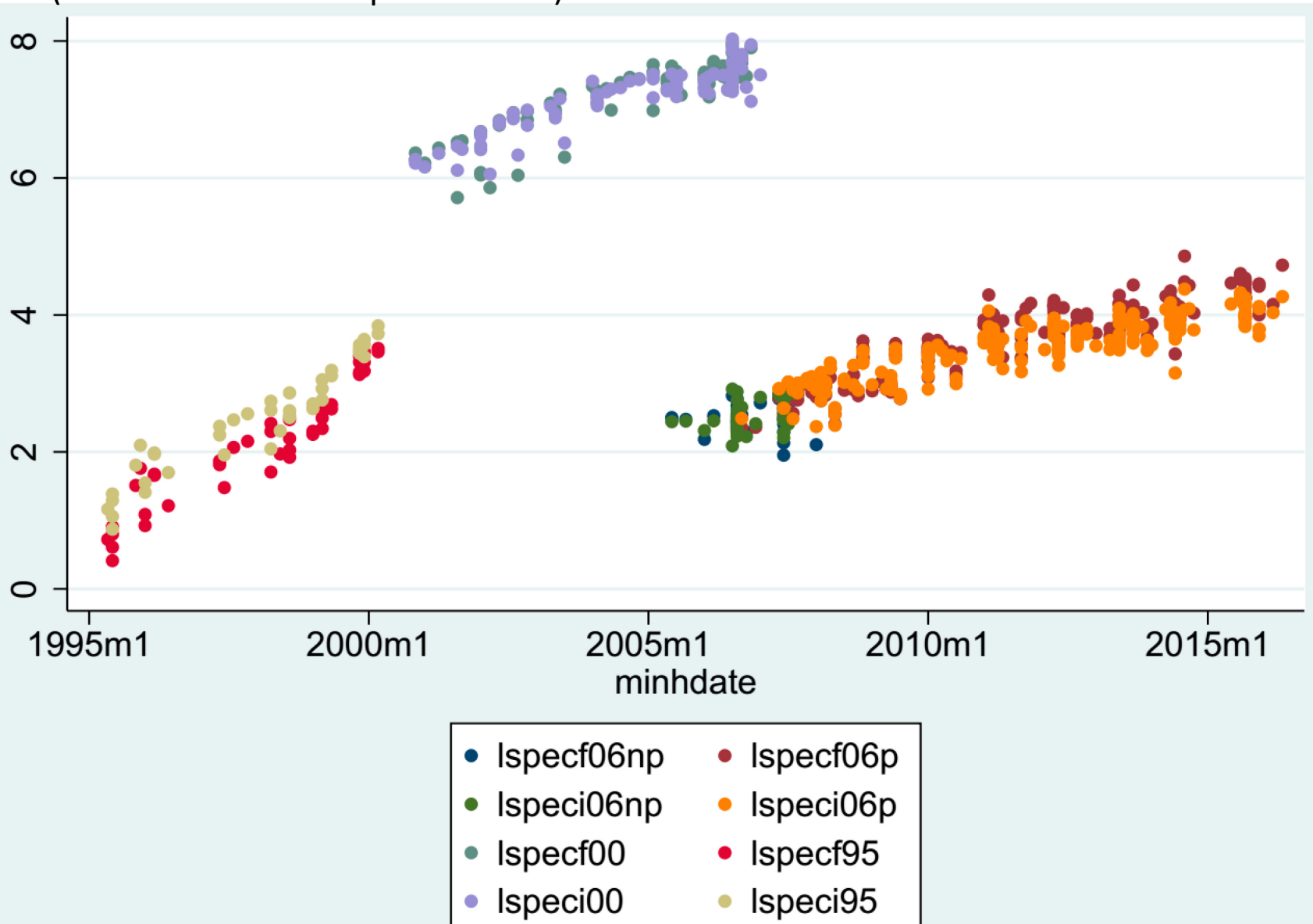
Evidence from Quality Indexes

CPU Quality Metrics?

- Estimation of time trend in SPEC CPU benchmarks shows continued decline in performance improvement rate after 2006

Desktop CPU Quality Metrics

Ln(Median Intel Desktop CPU Score)



SPEC CPU Benchmark	Coef. CAGR	Robust Std. Err.
1995m5-2000m3		
int95	.5826577	.0175146
fp95	.6397016	.0231907
int95_rate	.6241582	.0273672
fp95_rate	.7227752	.0331
2000m11-2004m11		
int2000	.3304092	.0173773
fp2000	.3429411	.023522
int2000_rate	.4697731	.0512966
fp2000_rate	.3989549	.0351676
2005m2-2007m1		
int2000	.3222474	.016442
fp2000	.3365855	.022279
int2000_rate	.4650892	.0475414
fp2000_rate	.3986346	.032545
2005m6-2012m11		
int2006	.1709304	.0069587
fp2006	.2467286	.0077563
int2006_rate	.2472256	.013015
fp2006_rate	.2537211	.0101781
2013m1-2016m5		
int2006	.1687175	.0064265
fp2006	.2414989	.0070952
int2006_rate	.2417978	.0119286
fp2006_rate	.2480768	.0093352

Conclusion

- Clear evidence of slowdown in price declines, cost declines, quality improvement for high volume chip types
- Classical Moore's Law scaling on last legs
- This is NOT the end of innovation in semiconductors
  - Just an important general purpose manufacturing technology dynamic within semiconductors, delivered a 20-30% annual cost decline for 40 years
  - Current focus in memory chips is 3-D device manufacturing using current technology nodes
  - Focus in mobile chips is clever software/firmware to reduce power use, better device integration
- Can current Moore's Law (smaller, but not necessarily cheaper) last ~~forever~~ another decade?