Moore's Law and the Economics of Leading Edge Semiconductors

Kenneth Flamm
Technology and Public Policy Program
Lyndon B. Johnson School of Public Affairs
University of Texas at Austin
kflamm@mail.utexas.edu

1

Low Tech Paper on a High Tech Subject

- To what extent is acceleration in pace of introducing new manufacturing technology linked to accelerating declines in semiconductor prices in late 1990s?
- What are implications of answer to this question?

2

Objectives

- Moore's Law today
  - "bumper sticker" for technology advance makes smaller, faster, cheaper
  - Smaller and faster not inevitably cheaper
- How does the technological smaller, faster, translate into the economic cheaper?
- Stylized Facts
- Link Technological Change to Economic consequences
- Embed "Moore's Law" In Economically Significant Framework
- Relate Stylized Facts to Institutional Changes in Industry
- Relate Framework to Recent Economics of Industry

3

Why Do We Care?

- Now largest U.S. manufacturing industry
  - Measured by value added
- One 4-digit manufacturing industry now almost 1% U.S. GDP
- Most important input to other manufacturing industries we care a lot about
  - Computers, communications
    - 40-60% of change in computer price
    - 15-30% of change in LAN hware price
    - Aizcorbe, Flamm, Kurshid (2002)
- Big impact on GDP, productivity growth
- See Jorgenson AEA 2001 Presidential Address

4

Globalized Semi Industry Challenges Economic Analysis

- Completely international production system
  - Different steps in process– Design, wafer fabrication, assembly, test- done at diverse locations
- Product mix varies greatly across countries
  - Example: US → 37%+ microprocessor IC share in 2002 vs. 20% globally
- Semi input mix varies greatly across user industries
- PPI's and appropriate input price indexes likely to diverge significantly
- Little reliable/useful data collected by govt.
- Industry-collected data is at company level, region of sale, no info on production region

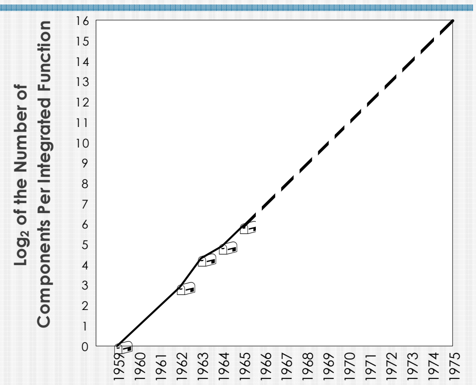
5

Moore's Law

- In the beginning: the original law
  - 2x devices/chip every 12 months
  - ca. 1965
  - Not about new technology originally–
    - About limits of existing tech vintage ("node")
    - chips the size of snack pizzas
- Moore rev.2
  - 2x devices/chip every 24 months
  - ca. 1975
- Split the difference became industry folklore
  - 2x devices/chip every 18 months, Moore rev 2.5
  - Unofficial schedule for introduction of new technology
- Self-fulfilling prophecy?
  - "it happened because everyone believed it was going to happen"
  - The receding brick wall

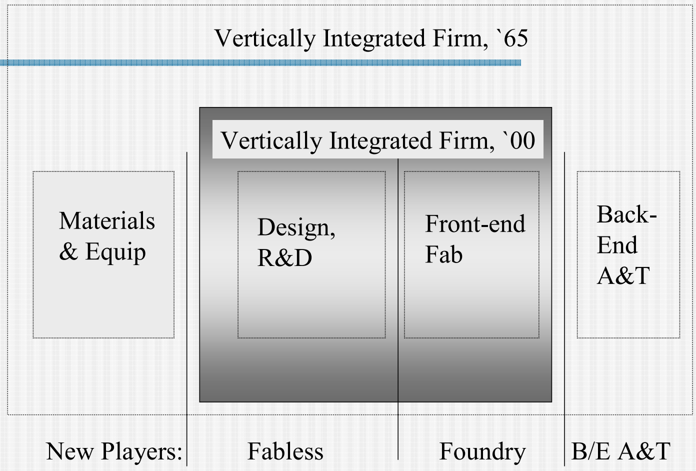
6

Figure 1
The Original "Moore's Law" Plot
From Electronics April 1965



7

Semi industry was also changing as Moore watched
…specialization created coordination issues



8

With Specialization Came New Coordination Issues

- Different pieces of increasingly complex technology now coming from proliferating numbers and types of vendors
- Only very largest leader firms attempted to coordinate next-gen tech internally
  - High cost
  - Accepted substantial spillovers to others
- “Moore’s Law” the de facto benchmark
  - Competitive target for device, equip producers

9

Focus on products making use of leading edge manufacturing technology

- Leading edge memory-DRAMs
  - DRAMs the canonical leading edge product through the 1990s
  - 70+% leading edge technology
- Leading edge logic-Microprocessors
  - Microprocessors increasingly the pacing product for leading edge semiconductor technology
  - 90% leading edge technology
- Story describes leading edge technology
  - Not everything is leading edge
  - In fact overall about 25 % processed wafers “leading edge” in 1999

10

What do semiconductor technologists model?

$$\text{Mfg Cost/device} = \frac{\underbrace{\frac{\$\ \text{processing cost}}{\text{area silicon}}}_{c}\ \underbrace{\text{Area/chip}}_{A}}{\underbrace{\text{Devices/chip}}_{d}}$$

c, A, and d well understood concepts

device = transistor/bit/etc., or equivalent

Trends frequently discussed among engineers, business strategists

11

Alternative formulation

$$\text{Mfg Cost/device} = \frac{\frac{\$\ \text{processing cost}}{\text{area silicon}}\ (c)}{\text{devices/Area silicon}\ (t)}$$

t is what is improved with innovation in Semiconductor lithography

t = d / A, so Moore’s law (d) + A or t determines t or A

new t = “technology node” (an approximation)

feature size reduced by 30%, device area by 50%
With new node

12

What do economists measure?

- Quality-adjusted price indexes
  - Hedonic methods
  - Matched model approximations
- Link to economics of technology:
  - Nomenclature issues:
    - Quality = functionality as valued by consumer
- Quality-adjusted price indexes (I) measure chip price/chip quality
  - P nominal chip price
  - f chip quality/functionality
    - May be determined implicitly
  - I=P/f quality-adjusted price index
    - f=P/I defines functionality/quality/"real" output per chip

13

Economists – technologists

$$\underbrace{\text{Price/device}}_{p} = \underbrace{\text{markup}}_{m} \times \frac{\underbrace{\frac{\text{\$ processing cost}}{\text{area silicon}}}_{c} \; \underbrace{\text{Area/chip}}_{A}}{\underbrace{\text{Devices/chip}}_{d}}$$

Comparing year t with year t+N, then, can decompose

$$\dot{p} = \dot{m} + \dot{c} + \dot{A} - \dot{d}$$

where $\dot{x}$ denotes $\ln(x_{t+N}/x_t)/N$, and compound annual growth rate (CAGR) for x over N yrs is defined as $e^{\dot{x}}-1$. For small CAGR only, CAGR = $\dot{x}$.

Moore's Law is about **d** only. Data on m, c, & A are needed to say something about costs and prices.
With everything else constant, faster Moore's law (bigger change in d) means faster decline in P.
But EVERYTHING ELSE NOT CONSTANT!

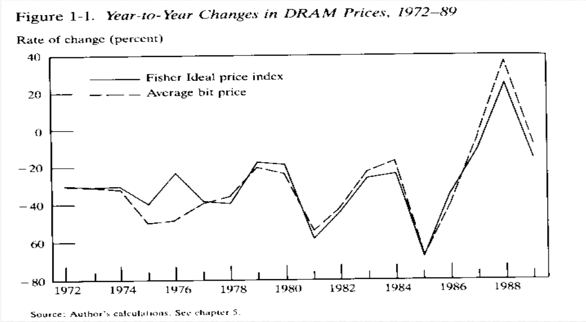
14

But economists don't measure Price/Device

- Economists measure **quality-adjusted** price I
- Price/device=I * f
- f is functionality/device
  - Functionality per bit or transistor can vary even within a single type of chip
    - Example: speed "binning" of DRAMs or microprocessors
    - "bell curve" for functional clock rates
    - some chips faster, sold for more, some chips slower, sold for less
    - Faster chips more valuable to consumers
- Would expect consumer valuation of chip functionality generally to not be exactly proportional to number of devices on chip
  - But would expect to observe positive relationship
- And some reason to believe that variation in f over time quite small relative to variation in I for some chips
  - Price/bit a tolerable long run approximation to I for DRAMs

15

Price per bit vs. Price Index in DRAMs



16

Restating economist-technologist interface

$$I * f = \underset{\text{markup}}{m} \times \frac{\dfrac{\$\ \text{processing cost}\ (c)}{\text{area silicon}} \quad \text{Area/chip}\ (A)}{\text{Devices/chip}\ (d)}$$

Comparing year t with year t+N, then, can decompose

$$\dot{I} = \dot{m} + \dot{c} + \dot{A} - \dot{d} - \dot{f}$$

where $\dot{x}$ denotes $\ln(x_{t+N}/x_t)/N$, and compound annual growth rate (CAGR) for x over N yrs is defined as $e^{\dot{x}}-1$. For small CAGR only, CAGR = $\dot{x}$.

So if f doesn't change much ($\dot{f} = 0$), we can think of quality-adjusted price index I and price/device approximations to each other. If not, we can estimate f as $\dot{f} = \dot{m} + \dot{c} + \dot{A} - \dot{d} - \dot{I}$

17

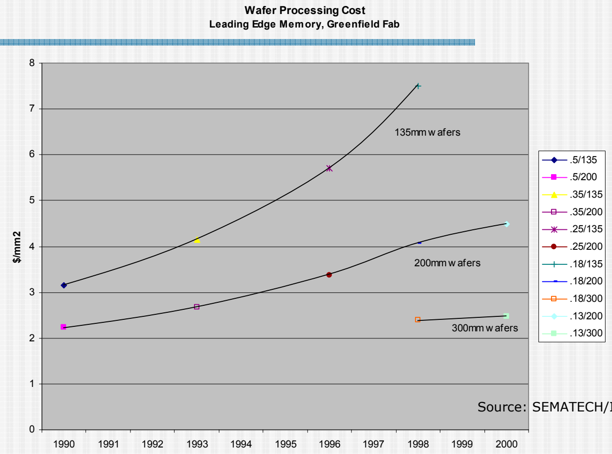
Leading Edge Memory Stylized Facts ➔ Data

- Devices per chip
  - Moore ver 2.5, 4x every 3 years
- Area per chip
  - New "technology node" every 3 years
  - Lithography advance alone ➔.5X area per chip feature
  - t-=2
  - Would predict Area/chip 2X every 3 years
  - Return to this in a moment
- **Processing Cost/Yielded Si Area**
  - i.e., per area of "good" chips
  - Roughly constant over time
  - Slow increase in unyielded cost offset by slow improvement in yields

- d
  - $\dot{d}$ = + 46%
  - CAGR = +59%
- A
  - $\dot{A}$ = + 23%
  - CAGR = 26%
- c
  - $\dot{c}$ = 0 %
  - CAGR = 0%

18

Leading Edge Memory Costs



Source: SEMATECH/ISMI

19

Assumptions

- “Long-run” view, ignore inessential issues
  - Learning economies, improvements in yields (more defect-free chips per wafer) within technology node (vintage)
  - Assembly and test economics not driving cost structure and prices
  - Assume markup constant in L-R for the moment
    - $\dot{m}=0$
- With, stylized facts from previous slide imply
    - CAGR = -21%
    - $\dot{P} = -23\%$

20

An Economist’s Default Corollary to Moore’s Law:

Moore’s Law + constant wafer processing cost + new technology node every 3 years

=

-21 %   CADR

21

Historical Reality at the Leading Edge

Decline Rates in Price-Performance

		Percent/Year
Microprocessors,	1975-85	-37.5
Hedonic Index	1985-94	-26.7
DRAM Memory,	1975-85	-40.4
Fisher Matched Model	1985-94	-19.9
DRAMs, Fisher Matched Model, Quarterly Data		
	91:2-95:4	-11.9
	95:4-98:4	-64.0
Intel Microprocessors, Fisher Matched Model, Quarterly Data		
	93:1-95:4	-47.0
	95:4-99:4	-61.6

Generally exceeded prediction!
Slowed down over time, then speeded up in mid-90s 22

The Ingenuity (DRAM) Corollary:

- Instead of doubling chip size, use ingenuity to increase it some factor < 2 times every 3 years
  - real recent example (DRAMs), 1.37
    - CAGR for A = 11%
  - 3-D device structures
- Implications of ingenuity
  - CADR = -30%, matches recent trend for DRAM
  - for DRAMs, in 70s and 80s, CADR more like -37%
    - wafer processing cost may also have fallen
    - Japan/VLSI project, competition impact?
- Another recent example is ASICs, shift toward more rapid leading edge technology adoption
  - transitory impact on CADR

23

Tinkering with Moore's Law: The Technological Acceleration (SEMATECH Roadmap) Corollary

- Suppose new technology node every 2 years instead of 3
  - Industry coordinated push through SEMATECH in 1990s: SEMATECH II
  - Later institutionalized in national, later international "roadmap" process
  - Competitive pressures also pushed
- New default (2X chip size)
  CADR = -29%
- New DRAM corollary (1.37X chip size)
  CADR = -41%
- Constant chip size (1X chip size)
  CADR = -50%
- Actual P improvement '95+ closer to -60%+!

24

Simulated Scenarios

		No of Years Between Nodes	Assumptions All Values in Period 0 = 1 m	c/y	t	Components of t d	A	Resulting p
3-yr node, 3-yr Moore's Law		**3**	**1**	**1**	**2**	**4**	**2**	**0.50**
no ingenuity	differences of logs		0	0	0.69	1.39	0.69	-0.69
	CAGR/CADR				25.99%	58.74%	25.99%	-20.63%
3-yr node, 3-yr Moore's Law		**3**	**1**	**1**	**2.86**	**4**	**1.4**	**0.35**
historical DRAM ingenuity	differences of logs		0	0	1.05	1.39	0.34	-1.05
	CAGR/CADR				41.90%	58.74%	11.87%	-29.53%
2-yr node, 2-yr Moore's Law		**2**	**1**	**1**	**2**	**4**	**2**	**0.50**
no ingenuity	differences of logs		0	0	0.69	1.39	0.69	-0.69
	CAGR/CADR				41.42%	100.00%	41.42%	-29.29%
2-yr node, 2-yr Moore's Law		**2**	**1**	**1**	**2.86**	**4**	**1.4**	**0.35**
DRAM ingenuity over 2 yrs	differences of logs		0	0	1.05	1.39	0.34	-1.05
	CAGR/CADR				69.03%	100.00%	18.32%	-40.84%
2-yr node, 2-yr Moore's Law		**2**	**1**	**1**	**4**	**4**	**1**	**0.25**
constant die size	differences of logs		0	0	1.39	1.39	0.00	-1.39
	CAGR/CADR				100.00%	100.00%	0.00%	-50.00%
2-year node, 2-year Moore's Law		**2**	**1**	**1**	**2.53**	**4**	**1.58**	**0.40**
DRAM ingenuity over 3 years	differences of logs		0	0	0.93	1.39	0.46	-0.93
	CAGR/CADR				59.11%	100.00%	25.70%	-37.15%
2-yr node, 3-yr Moore's Law		**2**	**1**	**1**	**2.51**	**2.51**	**1**	**0.40**
constant die size	differences of logs		0	0	0.92	0.92	0.00	-0.92
	CAGR/CADR				58.43%	58.43%	0.00%	-36.88%

25

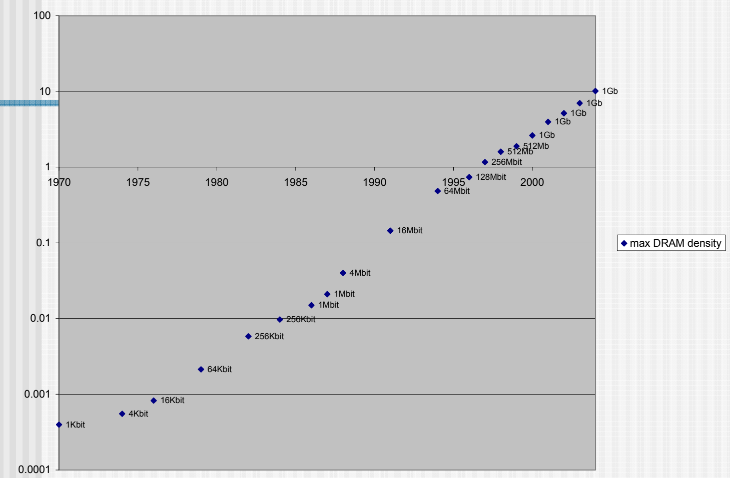
Nice Story, But Is It Really True?

- Bit density (t) actually grew more slowly in late 90s!

	cagr
1976-94	42.48%
1996-99	36.82%
2000-2004	40.14%
1976-2004	39.96%

26

Device Density for Drams with Highest Density



Source: ICKnowledge.

27

And Moore's Law may have picked up, but not by much in DRAMs

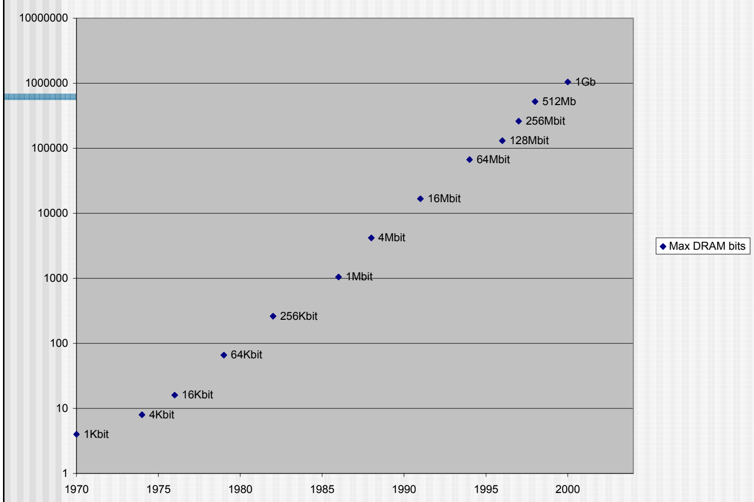
- Average was exactly Moore's Law over entire 1976-2000 period
- Significant questions about this data: are we picking up the most advanced players?

Growth in bits/chip, new chips

	cagr
1976-94	38.58%
1996-2000	68.18%
1976-2000	58.74%

28

Bits per DRAM, Newly Introduced DRAMs



29

Post '95 looks most like classic DRAM regime story

- Looking back at simulation table, like
  - Moore's Law classic
    - 3 years to quadruple bits
  - New tech intro classic
    - 3 years between nodes
  - Historical ingenuity
- Some acceleration in bits per chip
  - But offset by greater chip size growth
- Great story about tech intro acceleration, but doesn't seem to have much to do with DRAMs depicted in this graph
  - Greater decline rates in product price in late 1990s must have come from other factors
  - But, big caveat: quality of data unknown:
    - Are we looking at Samsung or Brand X?

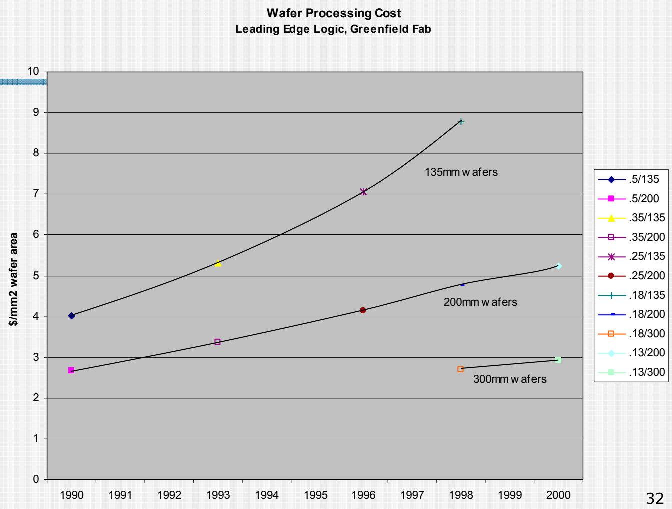
30

Intel Desktop Stylized Facts

- Constant L-R processing costs for silicon area
- Moore’s law (slightly below) for d
- Constant die size
- t CAGR 54-54%/yr

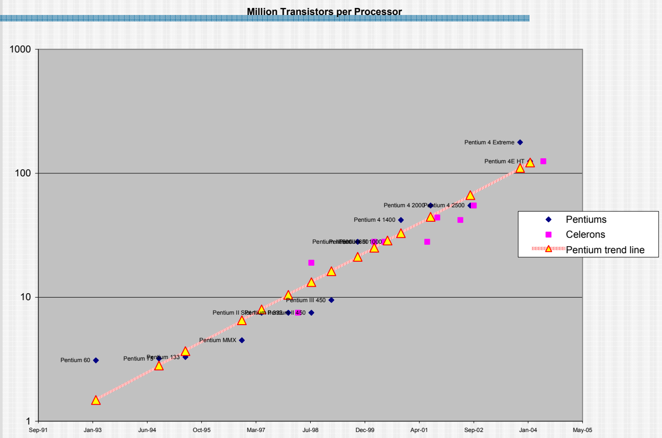
31

Microprocessor Wafer Processing Costs



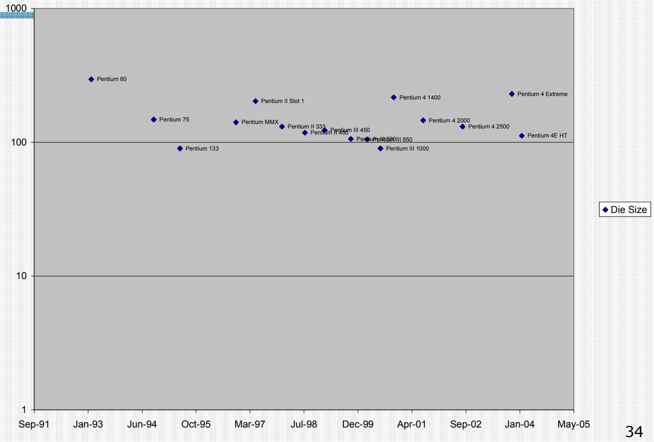
32

Pentium Moore’s Law



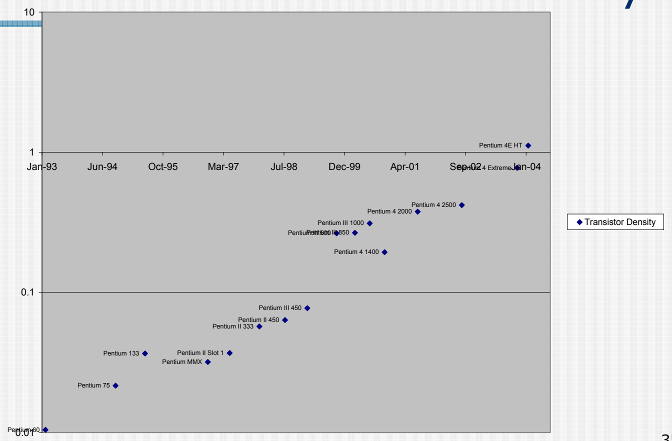
33

Pentium Die Size



34

Pentium Transistor Density



35

What do data tell us?

- Consistent with
  - 2-year node intro
  - historical (3-year) Moore d
  - constant die size
  - no ingenuity effect
- See prior simulations table

36

Intel Desktop Price Index

Intel Desktop Processor Price
Annualized Growth Rates

Period	Flamm (2004)	Aizcorbe, Corrado, Doms (2000)	
Feb. 1996-Feb. 1997	-52.66%	Q495-Q496	-55.23%
Feb. 1997-Feb. 1998	-68.92%	Q496-Q497	-62.90%
Feb. 1998-Jan. 1999	-68.43%	Q497-Q498	-69.20%
Jan. 1999-Jan. 2000	-64.80%	Q498-Q499	-76.40%
Jan. 2000-Jan. 2001	-69.23%		
Jan. 2001-Jan. 2002	-54.91%		
Jan. 2002-Jan. 2003	-53.86%		
Jan. 2003-May 2004	-59.15%		

Note recent decline in rate of decrease

37

Some Evidence of a Slowdown in Microprocessor Innovation

- Intel Desktop CPUs
- Estimated new hedonic price indexes
  - Price as function of very rich menu of desktop processor characteristics for different models
    - E.g., bus speeds, packaging, caches, in addition to usual processor speed characteristic
- Results consistent with previous work

38

More Evidence of Slowdown

Growth rate in transistors per chip could be interpreted as increase in rate of increase in transistors per chip in mid 90s, slowdown after 2000

SPEC benchmarks show declining performance improvement in Intel microprocessors

Annual growth rates in SPEC FP benchmark performance

June 95-March 2000	80%
March 99-August 2004	48%

Annual growth rates in SPEC Integer benchmark performance

June 95-March 2000	78%
March 99-August 2004	44%

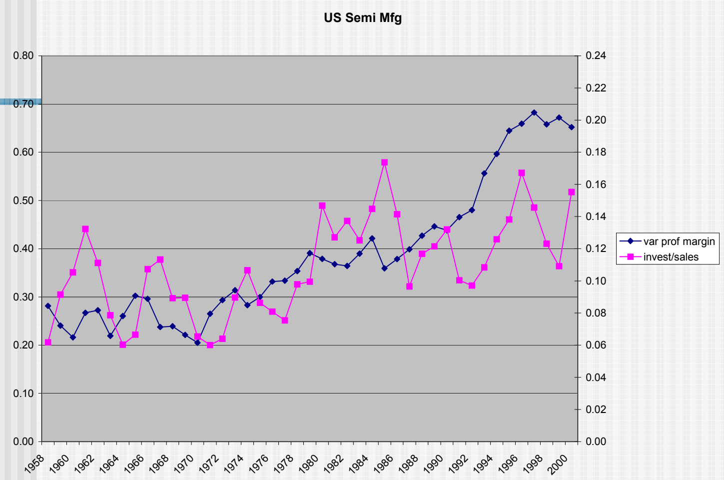
Note user functionality measure >> transistors/chip growth prior to 2000

39

Changes in Margins as Explanation for “Excess” Price Decline Rate in late ’90s?

- Some evidence suggesting lower margins in microprocessors
  - Aizcorbe (2002)
- Aggregate US data on semi margins not conclusive
  - Variable profit margin seems to have increased substantially!
    - From 40% in late ’80s to 65% in late ’90s
  - May reflect structural change in organization of US industry
    - Outsourcing manufacturing offshore
    - Integration of computer R&D into semi industry
  - May reflect changing product mix toward higher margins
    - Microprocessors 29% US IC shipments 1995, 37+% 2002
    - DRAMs 14% US IC shipments 1995, 7% 2001
- In DRAMs, microprocessors, data suggest trend toward **increasing** concentration in late ’90s
- Industry downturn in late ’90s may have lowered margins

40



41

Generalizing Moore

- Devices/chip:
  - Ver. 1
    - 2x every 12 months
      - CAGR: 100%
  - Ver. 2
    - 2x every 24 months
      - CAGR: 41%
  - Ver. Current
    - 2x every 18 months
      - CAGR: **59%**
- More Generally
  - Total quality/functionality/”real” semi output =total expenditure ÷ quality-adjusted price index
  - Multiply both sides by price index, divide by total chips
  - Implicit Index of “Functionality/quality” defined by quality-adjusted price index
    - Func/chip = Price/chip ÷ Price/quality
  - Can use to examine extent to which functionality/chip improvement tracked Moore transistor/chip improvements
  - Completely independent of earlier decompositions
    - Except for (possibly) using price index I

42

Interpreting Quality Improvement

- DRAM on average = Moore's law functionality/chip improvement
  - $/quality = approx. $/bit
- Microprocessors improve 2.5 X Moore's law prediction
  - In part, because microprocessors may have *become* "leading edge" product in 1990s
- ➔Microprocessor quality improvement dominated by factors above and beyond more transistors, qualitative innovation
  - $/quality << $/transistor

Improvement in Quality/chip

	CAGR		
	91-95	95-99	91-99
Microprocessors	112.2%	176.2%	142.1%
DRAM	39.7%	76.4%	57.0%

43

Implications of This Moore's Law Analysis

- Ultra-high rate of innovation in late 1990's extraordinary
- Transitory factors may have increased innovation above sustainable rates
  - Shortened product lives
  - Intensified competition
  - More rapid adoption of leading edge processes in range of products
- Microprocessors a leader in innovation
  - Quality >> manufacturing improvement alone
- Economic impacts may decline to lower but more sustainable rates
- International pacing process now in plac e
  - International roadmap successor to informal Moore's law benchmark
  - Hard to slow down!

44

Backup

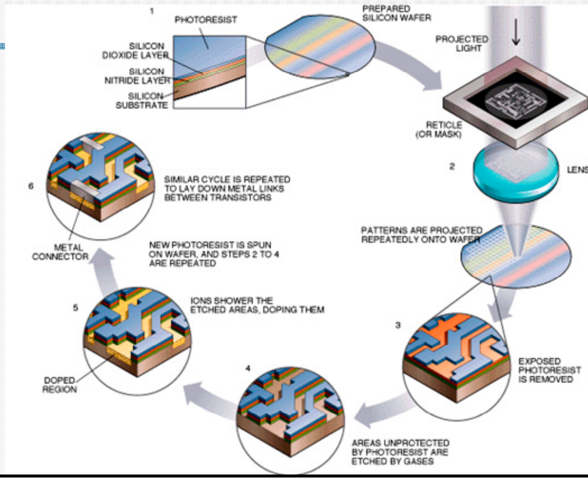
45

With Increased Size Came Increased Specialization

- ‘60s: backend A&T peels off
- ‘70s: materials and equipment separates
- ’80s: fabless design firms
- ’90s: designless fab firms- “foundry

46

IC Production



47