NORTHROP GRUMMAN
DEFINING THE FUTURE

Hard Disk Storage Deflation
Is There a Floor?

SCEA National Conference 2008
June 2008

Laura Friese
David Wiley
Allison Converse
Alisha Soles
Christopher Thomas

SCEA
The Society of Cost Estimating and Analysis

Hard Disk Deflation Background

- Storage Systems and Cost Estimates
  - Storage system costs are a major component of many automated information systems (AIS) cost estimates
  - Vendor quotes or specific storage specifications are sometimes available to estimate the cost of these storage solutions
  - Often, especially for significantly future storage cost projections, a $/volume of storage is applied to projected storage volumes
- Deflation Trends
  - Storage costs per unit volume historically deflate considerably over time
  - Using compounding deflation rates from year to year quickly leads to near zero $/volume costs for hard disk storage
    - To counter near-zero $/volume costs, cost estimators sometimes choose to use a “floor,” or point at which storage no longer deflates
  - For estimates that extend out past five to six years, the choice of the floor value becomes a significant input to the storage cost estimate

Overview and Objectives

NORTHROP GRUMMAN

- This paper explores whether or not these "floors" exist, by examining historical hard drive $/volume storage costs against current hard drive values
- Only desktop and laptop internal hard drives are examined
  - Data set spans 1996 to 2008 and was obtained from vendor print advertisements and online quotes
  - Personal computer hard drives were chosen as this data is readily available, spans long time periods, and has characteristics that are simple to normalize for
    - This includes size, interface format, and standard computer hard drive size of that time period
  - Also, personal computer hard drives are easily separated from the desktop or notebook on which they reside
    - With other storage devices, it can be difficult in practice to separate the actual hard disk from enclosures, racks, cases, and even servers in the quotation prices
    - It is important to separate out the hard drives from the rest of the equipment which may have different cost deflation properties
- This paper primarily examines $/volume costs
  - Future storage systems may require additional storage to accommodate increased systems requirements or file sizes
  - To correctly use $/volume factors to estimate storage cost requires capturing any future trends on file sizes or total storage required
- Part 1 will explore estimating hard disk storage deflation and potential floors when storage volumes have been correctly estimated
- Part 2 will explore hard disk storage trends, considering minimum and average hard disk sizes over time

LauraMFriese@ngc.com, March 2008

NORTHROP GRUMMAN

Part 1: Hard Drive Storage Deflation

If you have correctly estimated your storage volume needs, is there a floor to the deflation?

LauraMFriese@ngc.com, March 2008

Hard Disk Storage Cost Deflation Rates

NORTHROP GRUMMAN

- Many studies conclude that hard disk storage costs deflate over time*
  - 1999 study by Cormier and Blackburn showed the rate of deflation in the hardware cost (which includes storage) was 37% per year
    - Argued that this cost deflation is perfectly offset by technology inflation of 59% per year
  - 2000 study by Gilheany reported the rate of deflation for magnetic disk storage in 2000 to be 45%
  - 2003 study by Zaffos found a continued decline of 35-40% per year in hard disk system prices
  - Cost Deflation vs. Technology Inflation of RAID Storage Systems (Converse, Watkins, SCEA 2006) supports the claim that RAID storage costs deflate 30-40% per year

- Technology cost deflation is generally thought to be due to density increases*
  - Moore's Law on integrated circuits
    - Moore's statement in "Cramming More Components onto Integrated Circuits", Electronics Magazine, 19 April 1965
      - "The complexity for minimum component costs has increased at a rate of roughly a factor of two per year... Certainly over the short term this rate can be expected to continue, if not to increase. Over the longer term, the rate of increase is a bit more uncertain, although there is no reason to believe it will not remain nearly constant for at least 10 years. That means by 1975, the number of components per integrated circuit for minimum cost will be 65,000. I believe that such a large circuit can be built on a single wafer."
    - Popular Moore's law: The doubling of the number of transistors on integrated circuits every 18 months
  - Magnetic storage density studies
    - Kryder's law: A law specific to hard disk storage costs, Kryder states that the density of information on hard drives has doubled every 13 months
    - In 2000, Gilheany cited IBM as reporting in 1994 that there is a 60% increase in areal bit density per year, which translates to a 37.5% decrease in magnetic storage costs per year

* References from Cost Deflation vs. Technology Inflation of RAID Storage Systems (Converse, Watkins, SCEA 2006)

LauraMFriese@ngc.com, March 2008

Use of Floors with Hard Disk Deflation

NORTHROP GRUMMAN

- Using 30-40% deflation rates quickly leads to near zero $/volume costs for hard-drive storage
- Can storage ever be free?
  - Exponential curves never actually reach zero, just very close to zero values, so using these does not predict "free" storage
  - Floors are applied to this deflation to counter near-zero $/volume storage costs
  - Keep in mind that the hard disks are just one part of the overall storage solution
    - These disks do not exist independently of servers, racks, peripherals, and workstations which are not captured in these $/volume costs
    - If these costs are also estimated at a $/volume ratio, the total $/volume storage cost never gets to zero
    - These costs can also be viewed as a "floor" if they do not deflate over time, or deflate at lower rates
    - **This paper only focuses on hard disk cost trends, separating out the problem of hard disk deflation vs. other storage components' deflation**
  - If required volume also increases over time, these near-zero $/volume costs are applied to increasing volumes, so actual realized purchase cost is not necessarily very close to zero
    - **Part 1 of this paper only focuses on $/volume metrics, separating out the problem of $/volume costs vs. total storage purchase trends**

LauraMFriese@ngc.com, March 2008

Data Set Description

NORTHROP GRUMMAN

- Cost and size data on personal computers and personal computer hard drives obtained from advertisements and magazine reviews from back issues of PC Magazine, PC World, Computer Shopper, and current (Feb 2008) online manufacturer websites
- The majority of the data was collected from retailer lists of available drives or computers, so reflect a reasonable understanding of the smallest and largest popularly available drives at that time period
- More data was readily available from 1996 until 2003
  - After this time period, retailer advertisements did not generally contain long lists of available hard drives and computers along with prices
  - It is likely that the from 2003 on, the stores that used to advertise their prices for a variety of hard drives did so online, vice in print
  - Data from 2004-2007 was mostly obtained from magazine article reviews, vice one store's offerings
- Personal computer hard drives
  - Data includes only internal hard drives
  - Data collected includes storage size (in GB), price, and interface format
  - Data includes only one manufacturer's hard drives, though multiple interface types are included
    - SCSI, ATA, and SATA
  - 318 data points on internal drives from Feb 1996 to Feb 2008
- All prices inflated to CY08$
- For regression purposes, time was measured in years since 10/1/1995

Exponential Modeling and Choice of an Asymptote

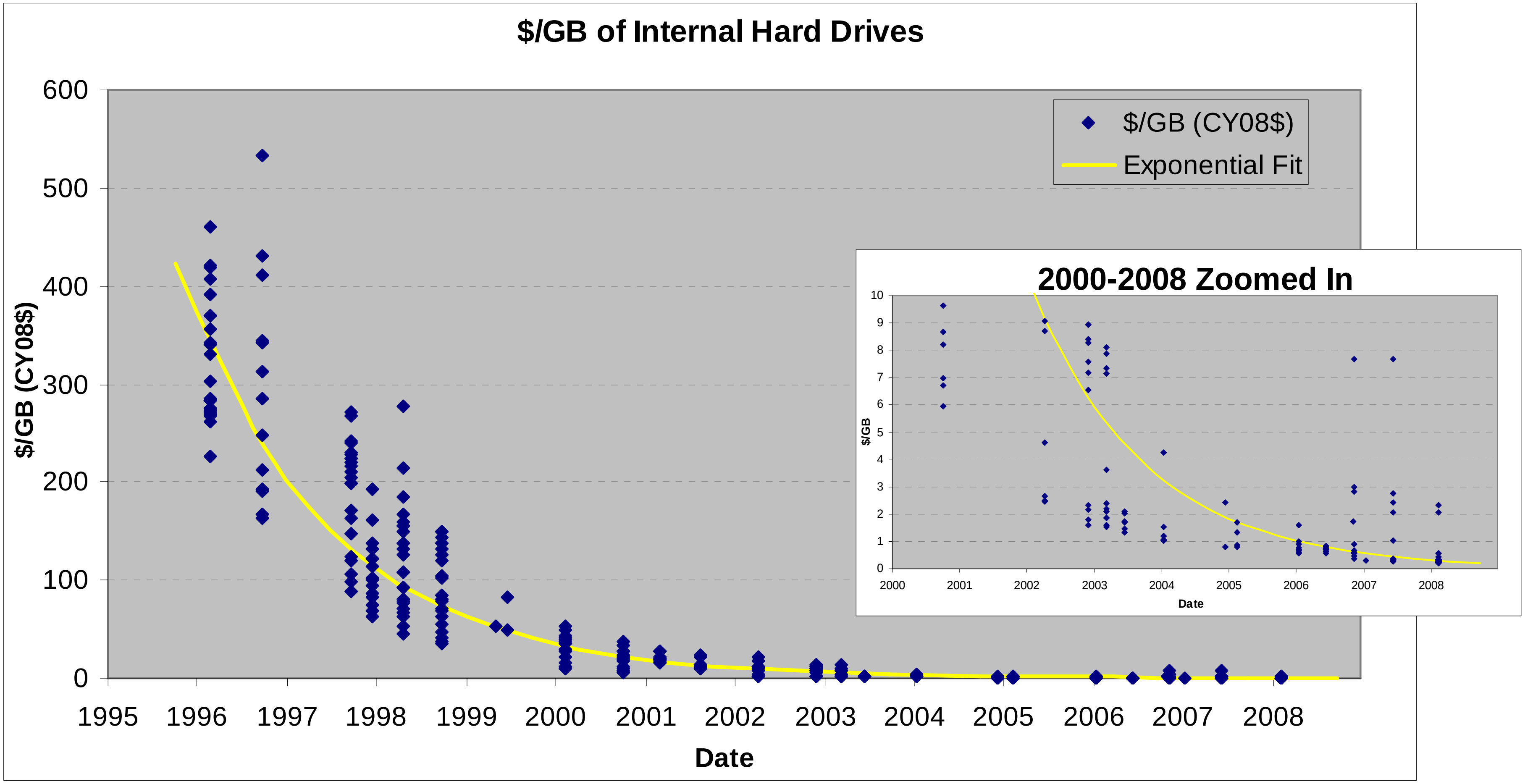
- With exponential modeling, the asymptote must be a given constant, not an independent variable
  - A modeler often already has an idea of what the asymptote is
    - For example, with Newton's law of cooling, a hot liquid cools exponentially to the room's ambient temperature
    - The room's temperature, which is known, is the asymptote
  - In other words, an estimator needs to guess at the asymptote value, and then model an exponential curve with that value
- For this data set, the following floors were considered
  - 10% of first year's minimum $/GB
  - Minimum $/GB in data set
  - Half of minimum $/GB in data set
  - Other potential floors were considered but ruled out because they were greater than points in the data set

LauraMFriese@ngc.com, March 2008

Exponential Regression of Full Data Set (No Floor)

NORTHROP GRUMMAN

- Exponential regression on full data set is significant (p-value < 0.05) and an R^2 of 0.94
  - The regression results indicate 44% deflation per year

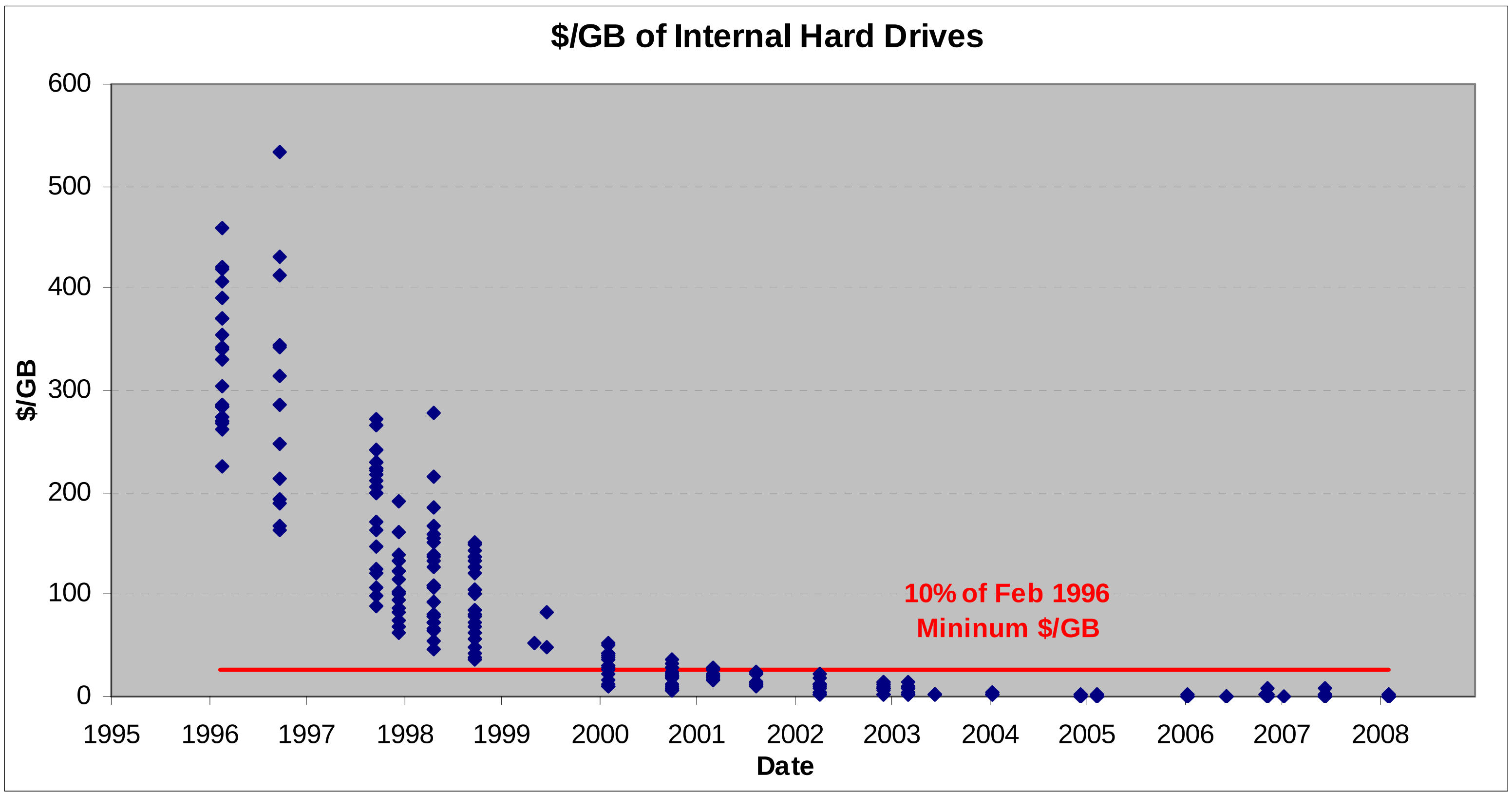


LauraMFriese@ngc.com, March 2008

Choice of Storage Cost Floor Using 1996 Data

NORTHROP GRUMMAN

- From the viewpoint of an estimator in 1996, with data available only from that year, what would be logical choices for the floor value?
- Estimator may choose a percentage of lowest $/GB in set as the potential floor
  - For example, choose 10% of minimum 1996 point, which is $22.6 $/GB
  - With 44% deflation applied to any 1996 starting data point, this floor is reached by 2002
- In full data set, data values below this are seen in 2000
  - All data points below this floor by 2002

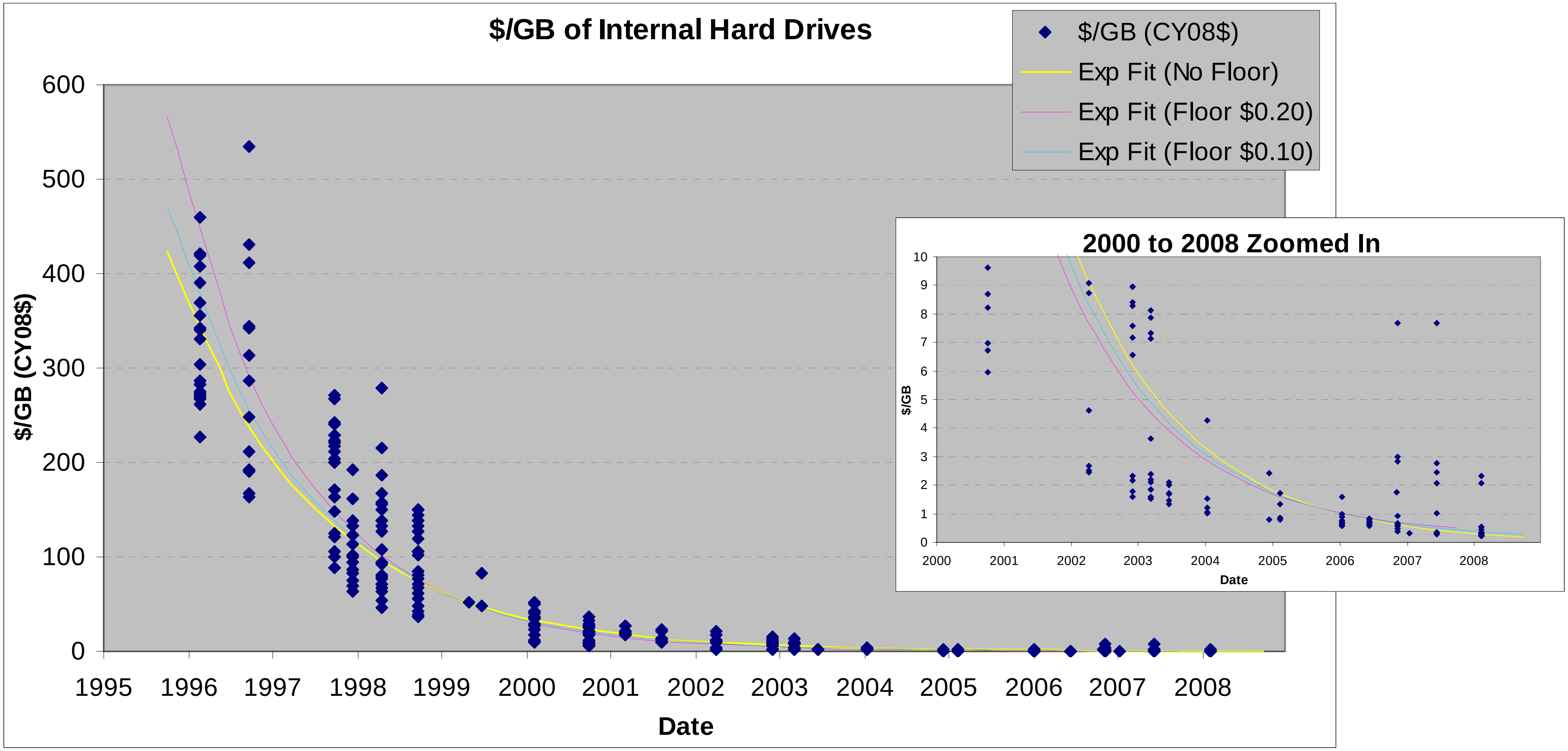


LauraMFriese@ngc.com, March 2008

Choice of Storage Floor Using Full Data Set: Minimum Point in Data Set

NORTHROP GRUMMAN

- Two other choices of floor tried are just below the minimum data point and 50% of the minimum data point
  - Minimum data point is \$0.21, so floors of \$0.20 and \$0.105 examined
- Both regressions significant, with R^2 values very close to the no floor regression's (0.93 and 0.94)
  - Since floor values are so small in comparison to full data set, use of these small floors are statistically equivalent to no floors
- Floor exponential regressions give slightly different regression equations, but predict deflation rates similar to no floor deflation (46-49% deflation)



Floor Regression Conclusions

NORTHROP GRUMMAN

- For personal computer hard drives, a floor chosen as a fairly small percentage (10%) of lowest data point in 1996 is not a good predictor of $/GB costs 8-13 years later
  - All data points below this floor by 2002
- Choices of floors based on minimum data point in set (minimum value and half of minimum value) are very similar statistically from a floor of zero
  - All exponential regressions significant with similar R^2 values (0.93-0.93)
  - Floor values very small in comparison to full data set
    - Data set's max value: $534/GB
    - Trial floors: $0.20/GB and $0.105/GB
  - These trials only show evidence of very small floors
    - For example, $0.21 (lowest point in data set) is 0.09% of minimum point in 1996, 0.05% of maximum data point in set
- Exponential regressions tried predict 44-49% deflation per year
- There is the argument that "technology jumps" in the future could counter this deflation
  - However, this data set included multiple interface standards (IDE/ATA to "Ultra ATA" to SATA and SCSI to Ultra SCSI), each which could be considered at least a small "technology jump"
- This data only goes back thirteen years, but significant floors would likely be seen in this time period if they existed from 1996 to 2008

NORTHROP GRUMMAN

Part 2: Explorations in Potential Causes of Floors

Are hard drive storage costs more constant than $/volume cost deflation implies?

LauraMFriese@ngc.com, March 2008

How Much Storage is Actually Needed?

NORTHROP GRUMMAN

- Do the minimum and average computer hard drive sizes increase over time?
- If purchased volume increases over time, and if $/GB decreases over time, do these two trends counter each other, causing a flat cost of one hard drive?
  - Do you pay the same amount for a hard drive each year?
- Would an estimator in 1996 estimate that a 2008 desktop would require a 80-500 GB hard drive, when the range of hard drives popularly available in 2006 was 0.34-1.6 GB?

LauraMFriese@ngc.com, March 2008

Historical Desktop and Laptop Hard Drive Sizes

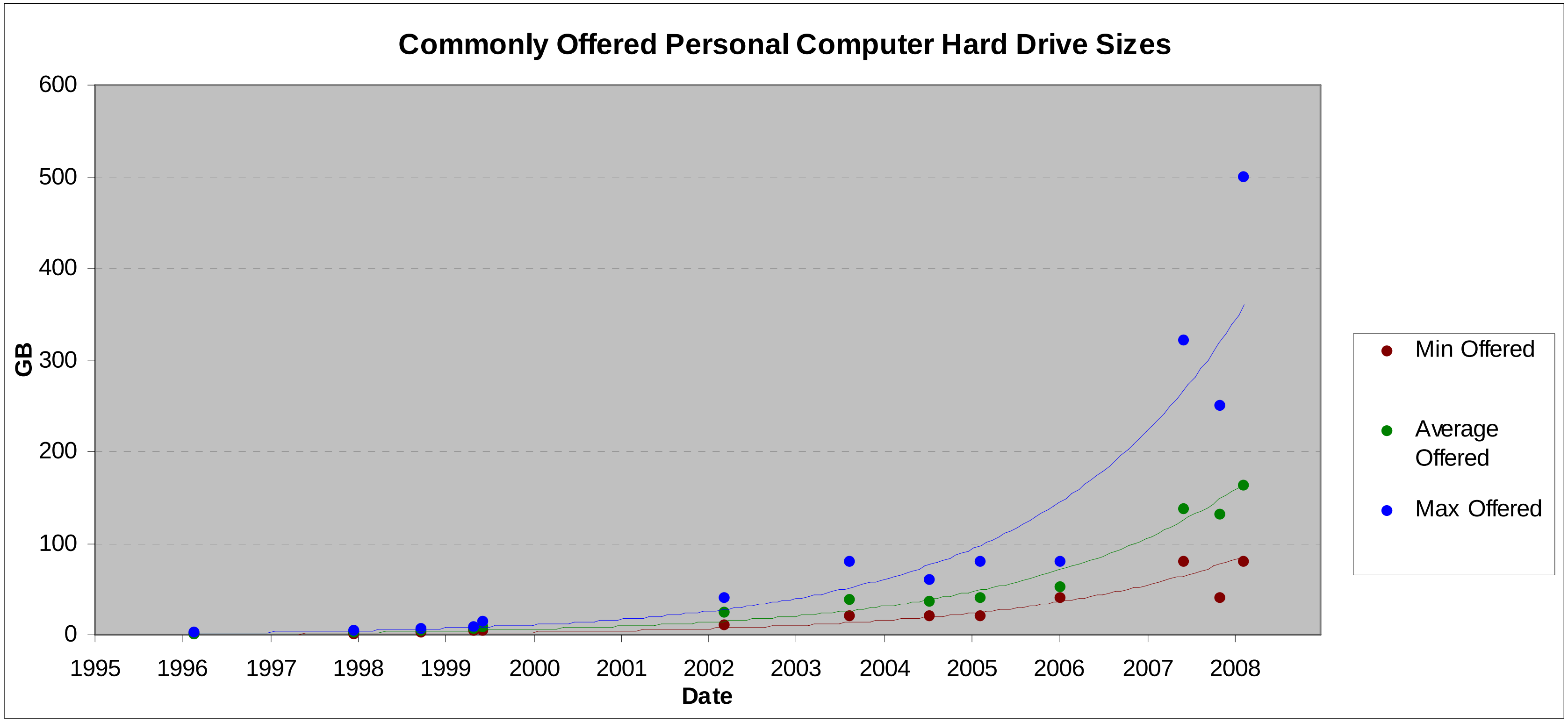
- In the same time period as the hard drive data, collected sizes of starting desktop and notebook hard drives
  - Data gathered in order to understand what the standard factory/starting sizes of computer hard drives were over time
- Data set displays three computer brands, and includes notebook and desktop computers
  - Data collected includes size of hard drive, manufacturer, and desktop vs. laptop
  - 200 data points from Feb 1996 to Feb 2008
  - All prices inflated to CY08$
- This data allows an understanding of what were the standard hard drive sizes over time
  - As with the other hard drive data set, there are probably smaller and larger hard drives offered as a base hard drive for the computers
  - These drive sizes were collected from long lists of available computers in the early years, and vendor ads in the out-years, so they do reflect *popularly* offered sizes

LauraMFriese@ngc.com, March 2008

Desktop and Laptop Hard Drive Sizes

NORTHROP GRUMMAN

- This graph shows minimum, average, and maximum personal computer drive sizes offered over the time period 1996 to 2008
- Each of these sets of data is best fit by an exponential curve (shown below), as expected by technology inflation research
- A potential cause of realized storage floors could be that the minimal or average drive cost stays flat over time
  - Disregard actual storage size and $/GB, but consider total cost per drive

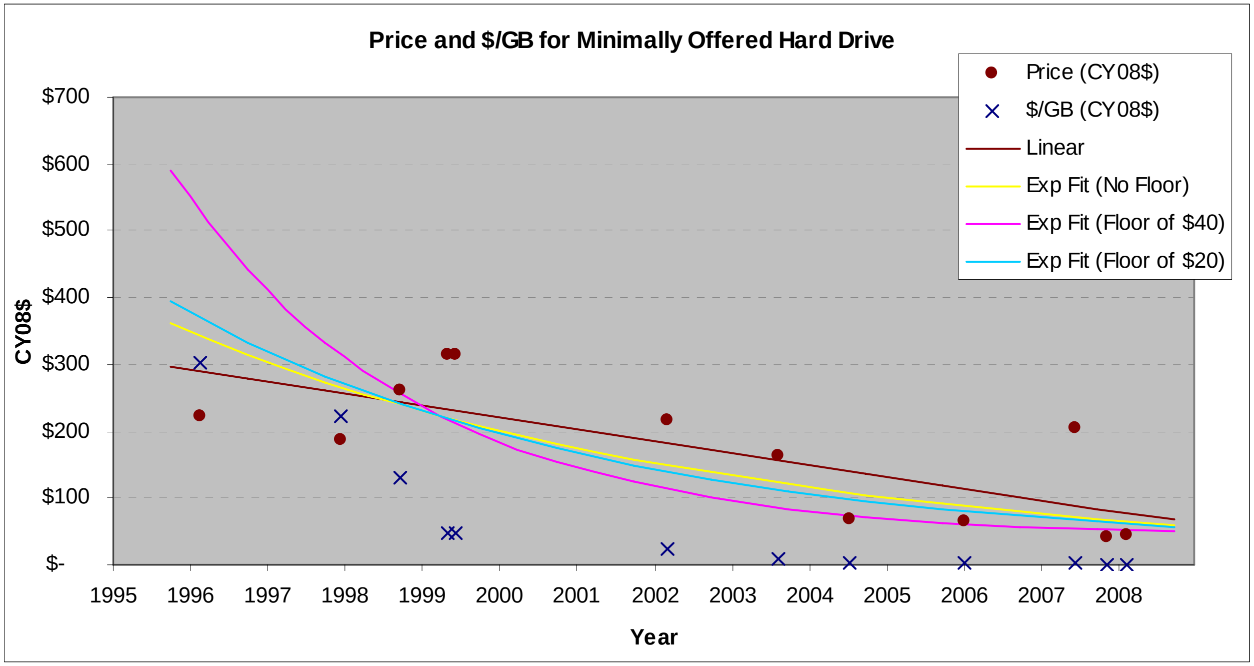


LauraMFriese@ngc.com, March 2008

Desktop and Laptop Hard Drive Sizes: Smallest Available

NORTHROP GRUMMAN

- The size of the smallest generally available hard drive increases over time
- How much does a “small” hard drive cost over time, understanding that the size of a “small” hard drive increases each year?
  - Is this a flat cost or does this decrease over time?
- For each of the minimum popularly offered computer hard drive sizes, the average cost of a hard drive of similar size, in that time period was examined
  - Corresponding $/GB is also shown as reference

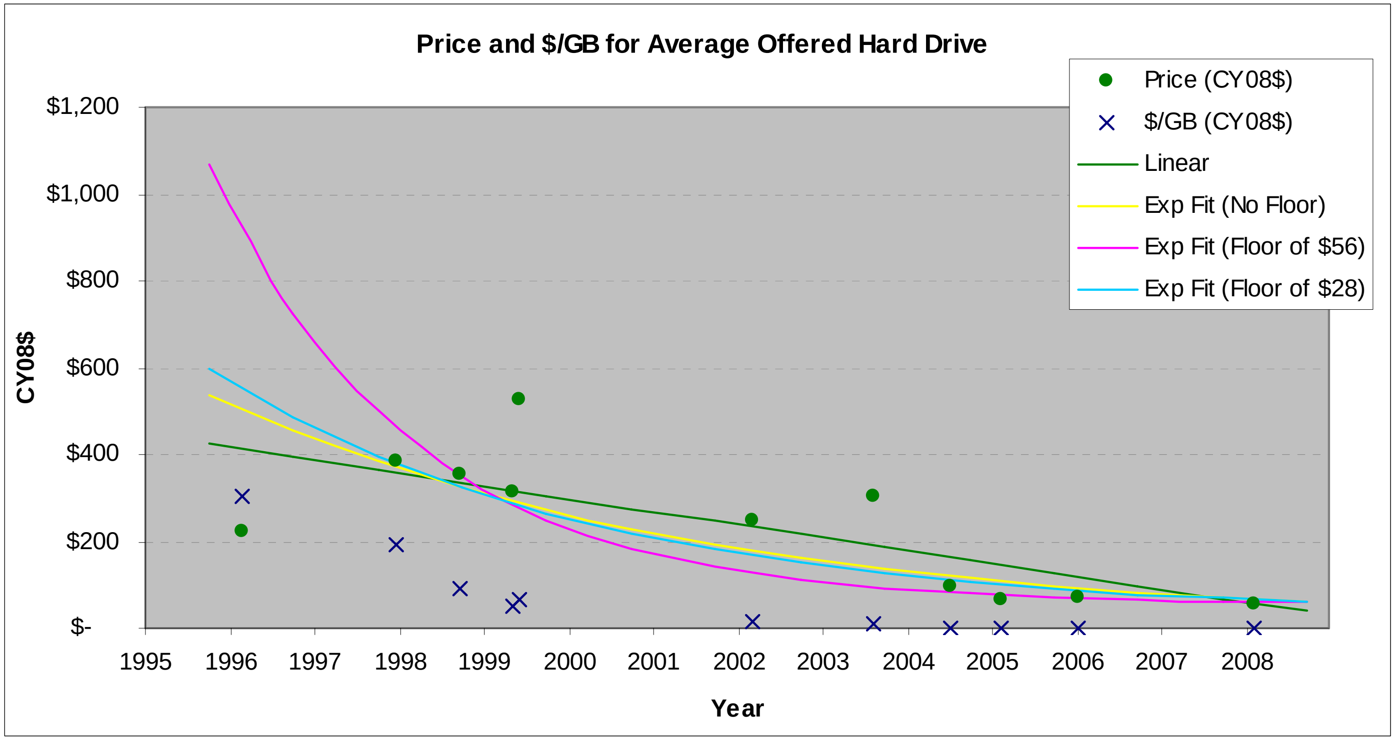


- For the price of a small hard drive, the following regressions were tried
  - Linear: Significant (p-value < 0.05), R^2 = 0.56,
  - Exponential with no floor: Significant (p-value < 0.05), R^2 = 0.58, annual deflation of 13%
  - Exponential with floor of minimum data point ($40): Significant ($p$-value < 0.05), R^2 = 0.53, annual deflation of 27%
  - Exponential with floor of half of minimum data point ($20):Significant ($p$-value < 0.05), R^2 = 0.58, annual deflation of 16%
- Best fits either an exponential curve with no floor or an exponential curve with a floor of $20
  - These predict deflation of 13-16% for minimum hard drive price

LauraMFriese@ngc.com, March 2008

Desktop and Laptop Hard Drive Sizes: Average Available

- A similar analysis was performed for the average hard drive size



- For the price of an average hard drive, the following regressions were tried
  - Linear: Significant (p-value < 0.05), R^2 = 0.54,
  - Exponential with no floor: Significant (p-value < 0.05), R^2 = 0.67, annual deflation of 16%
  - Exponential with floor of minimum data point ($40): Significant (p-value < 0.05), R^2 = 0.69, annual deflation of 34%
  - Exponential with floor of half of minimum data point ($20):Significant (p-value < 0.05), R^2 = 0.69, annual deflation of 20%
- Best fits with floor of $28-$56, though no floor is very close
  - These predict deflation of 20-34% of total average hard drive price

Part 2 Conclusions

- Answering the hard drive storage size questions previously posed:
  - Over time, the minimum and average hard drive size increases over time
    - $/GB still decreases exponentially over time
  - For a minimum popularly offered hard drive, best regressions are exponential with either no floor or a floor of $20
    - These predict deflation of 13-16% of minimum hard drive price
  - For an average popularly offered hard drive, best regressions are exponential with a floor of $28-56, though no floor is very close
    - These predict deflation of 20-34% for an average hard drive price
  - Since exponential deflation still does exist in this data, the $/GB is decreasing more quickly than increases in hard drive size
  - Deflation of overall minimum or average hard drive price is less steep than $/GB deflation
    - Indication that floors could exist to the overall minimum or average hard drive price
      - If these floors do exist, there may be a price that small or average hard drives do not go below, no matter how many GB are on the drive

Next Steps and Further Research

NORTHROP GRUMMAN

- The data set used does not reflect either enterprise or commercial storage systems, which a cost estimator more typically would be asked to estimate
  – Similar analyses could be done on enterprise level storage systems, being careful to separate out the disk costs from other storage components' costs
- Hard drives are just one part of the storage system
  – Deflation trends for the remainder of the storage systems should also be examined
- Increased hard drive density could cause increased power and cooling requirements
  – This could partially offset $/volume savings and these trends should be examined
- Other trends such as average enterprise file size, or document file size could be examined
  – This could cause increases in volume needed to store the same amount of material
- More data could be collected to understand the cost impacts of major technology jumps
  – For example, not just personal computer hard drive interface formats (SCSI, ATA, SATA), but more dramatic changes such as costs of tape vs. hard disk vs. optical storage

LauraMFriese@ngc.com, March 2008

Conclusions and Uses in Cost Modeling

- It is important to correctly estimate needed storage volumes
- For predicting the costs for recapitalized hardwareThis analysis could be considered for recapitalization hardware, consider applying less dramatic deflation rates, closer to the average hard drive price deflation rates
  - Will you be recapitalizing a 10GB hard drive in five years or a 100 GB hard drive? Have you correctly estimated the needed storage?
  - For hard drives, this would be mean applying 20-34% vice 44% deflation
- If you choose to use a floor to deflation, make sure the reasons are clearly stated
  - For example, if you feel that storage volumes have not been correctly estimated, and more storage will needed in the future
    - For storage hardware that is a recapitalized item, the recapitalized item may be larger in size than the original
  - For example, if the $/volume includes more than hard drives and you have reason to believe the non-drive hardware does not deflate as steeply
- If storage volumes have been correctly estimated, and you are just estimating the cost of personal computer hard drives, use of no floor or very minimal floors is reasonable based on trends in the last thirteen years

LauraMFriese@ngc.com, March 2008

References

- Converse and Watkins, *Cost Deflation vs. Technology Inflation of RAID Storage Systems*, SCEA 2006
- Cormier and Blackburn, *Technology Improvement and Cost Deflation*, Joint ISPA/SCEA International Conference 1999
- "Cramming More Components onto Integrated Circuits", Electronics Magazine, 19 April 1965
- Gilheany, *The Decline of Magnetic Disk Storage Cost Over the Next 25 Years*, March 2002, White paper available online at http://www.berghell.com/whitepapers/Storage%20Costs.pdf
- "Kryder's law", Scientific American Magazine, August 2005
- "Storage Systems", CFO Magazine, 17 November 2003 (Zaffos quote)

LauraMFriese@ngc.com, March 2008

Data Sources

NORTHROP GRUMMAN

- PC Magazine Issues:
  - 20 February 1996
  - 24 September 1996
  - 23 September 1997
  - 16 December 1997
  - 21 April 1998
  - 22 September 1998
  - 04 May 1999
  - 08 June 1999
  - 22 June 1999
  - 08 February 2000
  - 03 October 2000
  - 06 March 2001
  - 15 August 2001
  - 12 March 2002
  - 03 December 2002
  - 11 March 2003
  - 17 June 2003
  - 07 November 2006
  - 16 January 2007
- Computer Shopper Issues:
  - August 2003
  - July 2004
  - December 2004
  - February 2005
  - January 2006
  - June 2006
  - November 2006
  - June 2007
  - November 2007
- PC World Issues:
  - January 2004
- Seagate.com website (for February 2008 current prices)
- IBM.com website (for February 2008 current prices)

NORTHROP GRUMMAN

DEFINING THE FUTURE