Econ 140, Lecture for Week 3 (Lecture notes by Jeffrey Clemens)

Recap of Model of Physician Utility

Defining terms:

π = profit
b(q) = total patient benefit as function of quantity
MB(q) = marginal benefit of unit q (this is the derivative of b(q))
θ = weight on profit $(1 - \theta)$ is weight on patient benefit

So we can write:

$$\text{Utility} = \theta\pi + (1 - \theta)\, b(q)$$

What's profit? Profit depends on how the physician is paid. To capture the major ways in which physicians are paid, we'll allow for both a salary component and a fee-for-service payment.

α = salary payment
r = reimbursement rate. (This is part of the payment that shows up in our diagrams.)
c(q) = total cost as a function of quantity
MC(q) = marginal cost of unit q (this is the derivative of c(q))

So $\pi = \alpha + rq - c(q)$

$$\text{Utility} = \theta[\alpha + rq - c(q)] + (1 - \theta)\, b(q)$$

$$MU(Q) = \theta[r - MC(Q)] + (1 - \theta)MB(Q)$$

"General Solution:" Let's now take an abstract look "optimal" physician payment:

We'll derive the reimbursement rate that leads physicians to supply until MB = MC.

The physician optimizes by setting Q such that MU(Q) = 0

$$0 = \theta[r - MC(Q)] + (1 - \theta)MB(Q)$$

Knowing that physicians behave this way, the insurer wants to pick the r that leads physicians to provide Q* such that MB(Q*) = MC(Q*). We can find this by substituting MC = MB and solving for r.

> We're moving away from thinking about the shapes of the curves and focusing exclusively on where they intersect (MC = MB). This will tell us more about the optimal rate, but not about the costs of missing it.

Substituting MC(Q*) = MB(Q*) gives us:

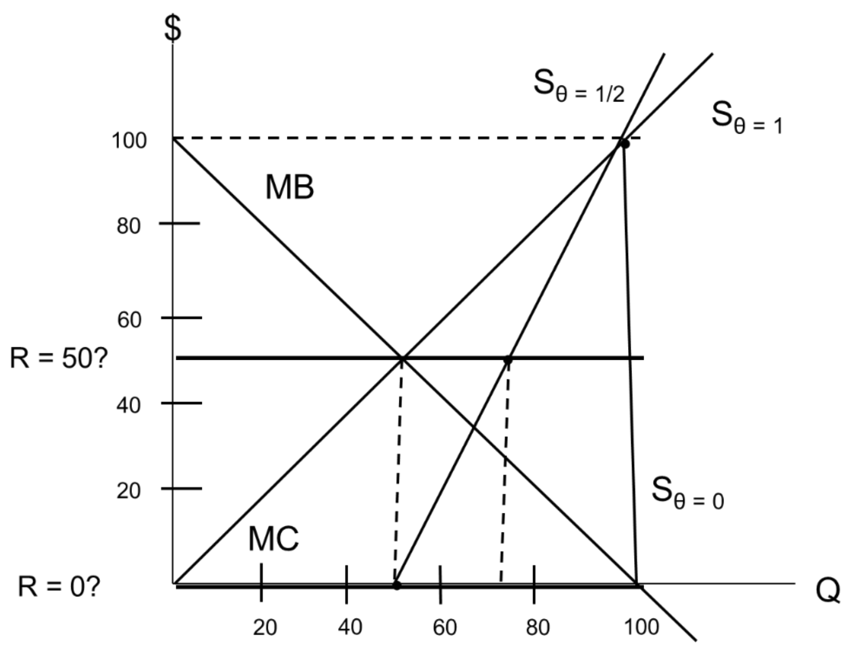
$$0 = \theta r^* - \theta MB(Q^*) + (1 - \theta)MB(Q^*)$$

Now solve for r*:

- ⇨ $\theta r^* = \theta MB(Q^*) - (1 - \theta)\ MB(Q^*)$ [add the terms involving MB to both sides]
- ⇨ $\theta r^* = \theta MB(Q^*) - MB(Q^*) + \theta MB(Q^*)$ [distribute the MB across 1 - theta]
- ⇨ $\theta r^* = 2\theta MB(Q^*) - MB(Q^*)$ [add the two theta MB terms]
- ⇨ $\theta r^* = MB(Q^*)[2\theta - 1]$ [factor out the MB]
- ⇨ $r^* = MB(Q^*)[2 - 1/\theta]$ [divide through by theta]

Thinking back to our initial diagram, we see two key pieces of intuition:

1) Optimal reimbursement is high when MB is high
2) Optimal reimbursement increases with θ. (Note that θ is in the denominator of a term that enters negatively on the right-hand side.)



Additional implications: for any given agency term, there's at most one optimal reimbursement rate:

1) If physicians vary, you can't get it right for all of them
2) For physicians who place particularly high weight on patient health, there's actually no rate that generates the desired quantity of care provision.

Billing Norms (how aggressively do physicians convert the care they provide into the billing codes they submit)

Let's add the complexity of billing norms to the equation.

Billing determines how care provided translates into measured/billed care. Suppose that for "real" quantity q, the physician submits bills for *aq*.

- Recall the example from last week's lecture. If you think your patient is "complicated" and submit the relevant billing code, the payment for an office visit adjusts upward by 50%!

Physicians' marginal utility from care provision reflects this adjustment to the financial benefit.

$$MU(Q) = \theta[ar - MC(Q)] + (1 - \theta)MB(Q)$$

Set equal to 0

$$\theta MC(Q) = \theta ar + (1 - \theta)MB(Q)$$

Substituting MC = MB gives us:

$$\theta MB(Q^*) = \theta ar^* + (1 - \theta)MB(Q^*)$$

Now solve for r*:

$$\theta ar^* = \theta MB(Q^*) - (1 - \theta)MB(Q^*)$$

$\Rightarrow$ $\theta ar^* = \theta MB(Q^*) - MB(Q^*) + \theta MB(Q^*)$

$\Rightarrow$ $\theta ar^* = 2\theta MB(Q^*) - MB(Q^*)$

$\Rightarrow$ $\theta ar^* = MB(Q^*)[2\theta - 1]$

$\Rightarrow$ $r^* = MB(Q^*)/a\ [2 - 1/\theta]$

Straightforward adaptation of previous solution. Just need to scale the reimbursement.

We again run into problems if physicians vary in how they bill.

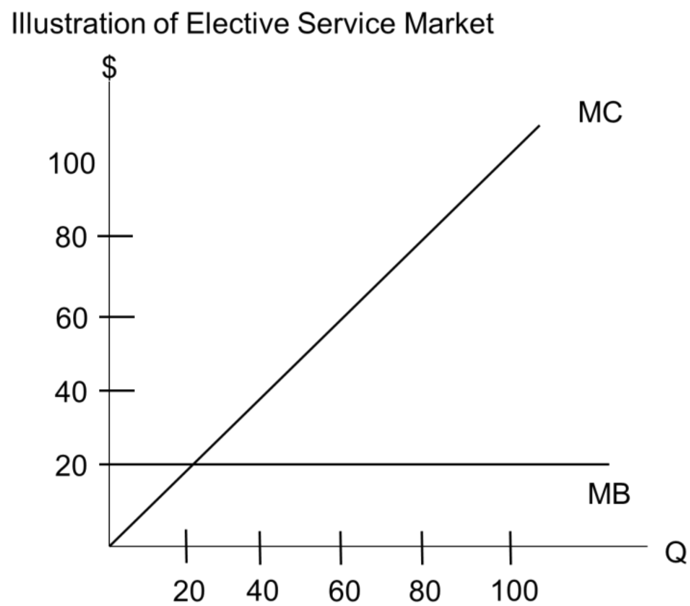
Variations across Services with Different Marginal Benefit Curves:

Now we're going to explore how things look when services have different benefit profiles.

Standard examples:

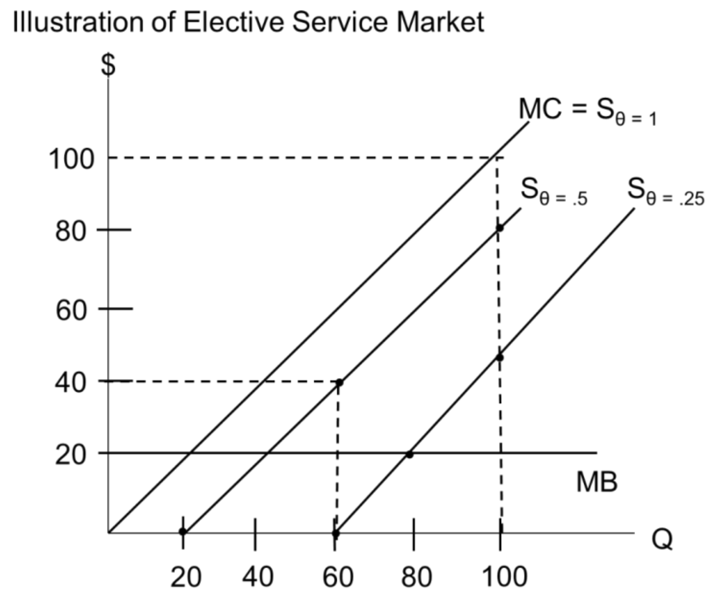
Elective services like diagnostic tests: flat benefit curve
Emergency-type services: High benefits for some, possible harm for others
Chemotherapy: High benefits for those with cancer, known harm for those without

Let's start with elective services. Assume MB = 20 for all. MC = Q



Let's go ahead and plug the MC and MB curves into the First Order Condition and solve for Q:

	$0 = \theta[r - MC(Q)] + (1 - \theta)MB(Q)$	[general first order condition]
⇨	$0 = \theta[r - Q] + (1 - \theta)[20]$	[substitute MC = Q and MB = 20]
⇨	$0 = \theta r - \theta Q + 20 - \theta 20$	[distribute both the theta and the 1 - theta]
⇨	$\theta Q = \theta r + 20 - \theta 20$	[add the Q theta term to both sides]
⇨	$Q = r + 20/\theta - 20$	[divide through be theta]



1) If the physician only cares about profit, then $\theta = 1$. This gives us
   a) Q = r - 20 + 20/1
   b) => Q = r (profit maximizer's supply curve traces out the MC curve as before)

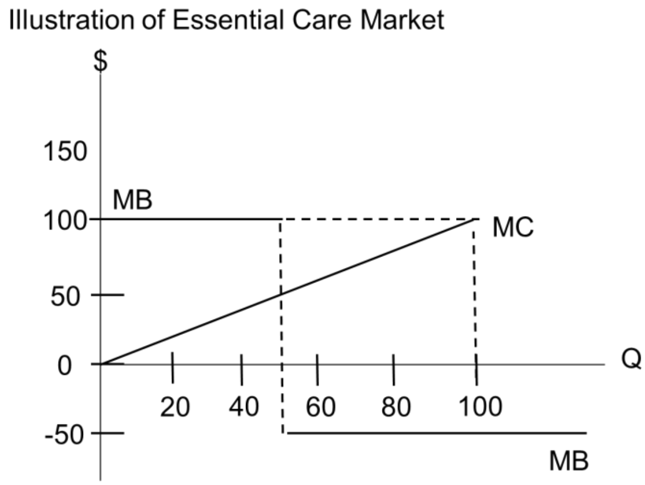
2) If the physician only cares about patient benefits, then $\theta = 0$. This gives us
   a. Q = r - 20 + 20/0
   b. => Q = ∞ (uh oh. MB never reaches 0)

3) What if the physician places weight of 0.5 on profit and 0.5 patient benefits?
   a. Q = r - 20 + 20/.5
   b. => Q = r - 20 + 40 = r + 20
   c. Let's plot some points on the supply curve
      i. With r = 0, Q = 20
      ii. With r = 40, Q = 60
      iii. With r = 80, Q = 100

4) What if the physician places a weight of 0.25 on profit and 0.75 patient benefits?
   d. Q = r - 20 + 20/.25
   e. => Q = r – 20 + 80 = r + 60
   f. Let's plot some points on the supply curve
      i. With r = 0, Q = 60
      ii. With r = 20, Q = 80
      iii. With r = 40, Q = 100

What's Happening Here: Recall from last time that when physicians have greater concern for patient benefit, the supply curve got steeper. We're now seeing that the steepening of the supply curve depends on the slope of the marginal benefit curve. If the marginal benefit curve is perfectly flat, then the supply curve doesn't actually get steeper at all as θ rises.

New Example: An Emergency or Essential Service
Now let's consider a service that has *high value for some of the population and negative for the remainder*. Assume MC = Q, that MB = 100 for first 50 units and MB = -50 thereafter [draw].



This gets tricky because there are two segments of the MB curve, which will mean two segments of the supply curve when physicians care about both profit and patient health.

We start with the segment of the supply curve for the first 50 units, where MB = 100

$0 = \theta[r - MC(Q)] + (1 - \theta)MB(Q)$ [first order condition]

⇨ $0 = \theta[r - Q] + (1 - \theta)[100]$ [substitute in MC = Q and MB = 100]

⇨ $0 = \theta r - \theta Q + 100 - \theta 100$ [distribute the theta and the 1 - theta]

⇨ $\theta Q = \theta r + 100 - \theta 100$ [add theta Q to both sides]

⇨ $Q = r - 100 + 100/\theta$ [divide through by theta]

Beyond 50 units, MB = -50. So we have a second segment of the supply curve described by:

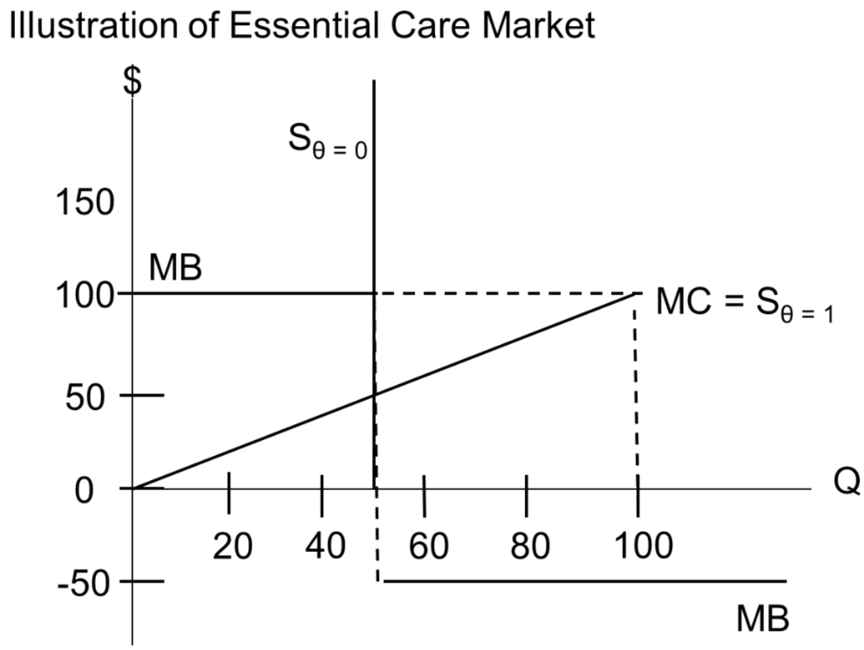
⇨ $0 = \theta[r - Q] + (1 - \theta)[-50]$ [substitute in MC = Q and MB = - 50]

⇨ $0 = \theta r - \theta Q - 50 + \theta 50$ [distribute the theta and the 1 - theta]

⇨ $\theta Q = \theta r - 50 + \theta 50$ [add theta Q to both sides]

⇨ $Q = r - 50/\theta + 50$ [divide through by theta]

Note that on both segments, we are left with Q = r = MC if $\theta = 1$

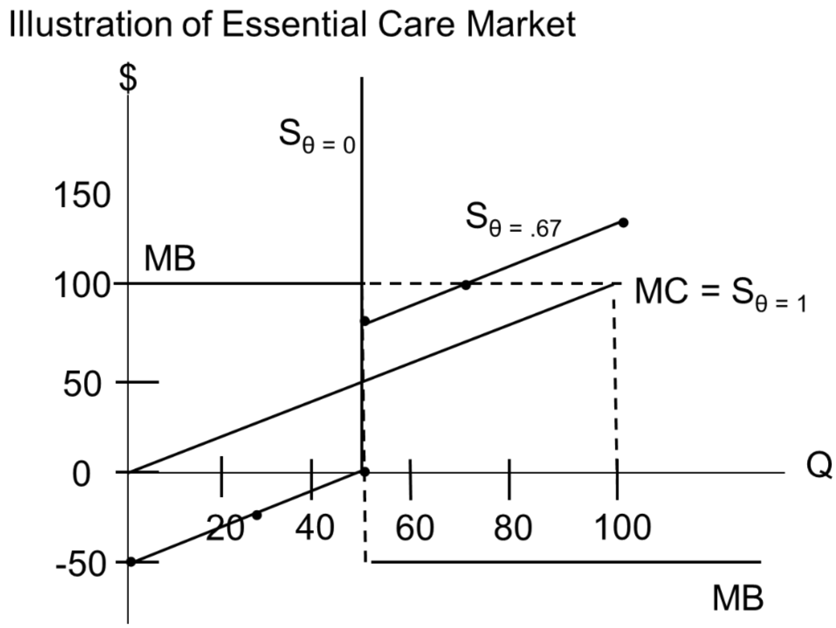


1) If the physician only cares about profit, then θ = 1. This gives us…
    a) Q = r regardless of which side of the discontinuity we are on. Note that this example only differs from the previous example in terms of the MB curve, so the profit maximizer looks exactly the same.

2) If the physician only cares about patient benefits, then θ = 0. This gives us
    a. Q = 50. Provide until MB drop below 0. (Note that the supply curve equation is undefined in this case. We need to use intuition and/or think through the graph)

3) What if the physician places weight of 0.67 on profit and 0.33 patient benefits? Let's plot some points on the first segment of the supply curve.
    a. Q = r – 100 + 100/θ
    b. Q = r - 100 + 100/.67
    c. => Q = r – 100 + 150 = r – 50
        i. With r = 0, Q = 50
        ii. With r = -25, Q = 25
        iii. With r = -50, Q = 0

    Now let's move to the second segment of the supply curve
    d. Q = r – 50/θ + 50
    e. Q = r - 50/.67 + 50
    f. => Q = r – 75 + 50 = r – 25
        iv. With r = 75, Q = 50
        v. With r = 100, Q = 75
        vi. With r = 125, Q = 100



Key Points:

What kind of physicians do we want?

1) In this setting neither pure profit maximizers nor pure health maximizers are ideal.
    a. In health care it can be important to have providers who would avoid providing profitable care if it is a cause of harm.
    b. Health maximizers will generally provide too much care (low MB)
        a) Some responsiveness to incentives allows the insurer to steer resources
2) If we could choose what kinds of physicians we want, we've seen that part of the answer is that we want *consistency* across physicians as much as we might want physicians of any particular type.
    a. This applies to both the way they submit bills and how much weight they place on patient benefits

Additional points:

3) There is more margin for error for services with steep benefit profiles ("essential" vs. "elective" services.
4) In general, the optimal reimbursement rate in this model is no more than your estimate of the marginal cost of the last unit of care you want the physician to provide.

Service Cost Structures
We'll now change gears to focus on cost structures rather than marginal benefit profiles.

We want to gain insight into how efficient payments will differ when we compare capital or technology intensive services to labor intensive services.

Real World Question: Why does the US have such extensive provision of MRIs and CT scans?

Answer: Medicare pays based on average costs. This is a bad for capital intensive services.
- Both the concept of average cost pricing and its implementation turn out to be relevant

The Basics: For all services, we can write that total cost = fixed cost + variable cost

$$TC = F + VC$$

To get an expression for average total cost we divide through by Q

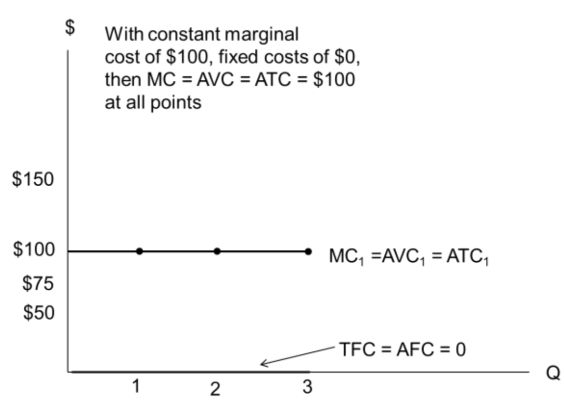
$$\frac{TC}{q} = \frac{F}{q} + \frac{VC}{q} \qquad Or \qquad ATC = \frac{F}{q} + AVC$$

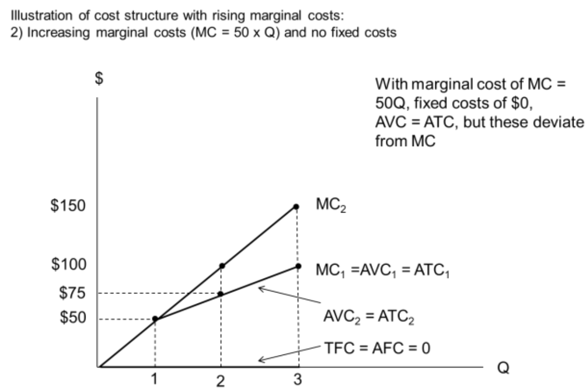
Think of F as a capital investment (e.g., MRI machine). For pure labor services, F = 0. For capital intensive services like MRIs, F can be quite high relative to the variable costs. (Also, note that with constant marginal costs, VC = q*MC, ans AVC = MC.)

Consider two services with no fixed costs (e.g,. a vaccination and a 30 minute office visit).
1) Constant marginal cost of $100.
2) MC = 50Q

The cost structures are illustrated in the diagrams below. Medicare would pay its estimate of ATC (average total cost), which in these cases equals AVC (average variable cost).

Illustration of cost structure for a hypothetical service:
1) Constant marginal costs and no fixed costs





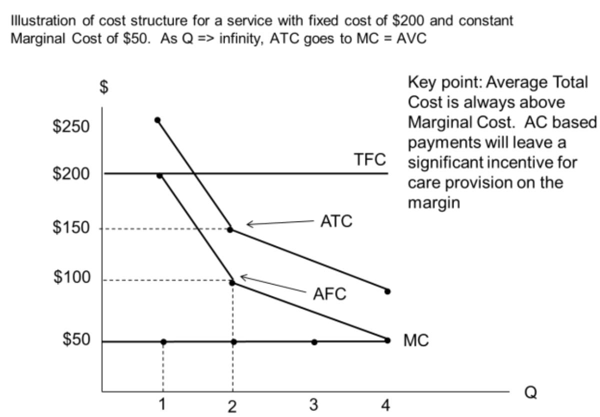
Key Points:

1) For the service with no fixed cost and constant marginal cost, we can see that average variable cost, average total cost, and marginal cost are always the same. This will also equal Medicare's payment.

2) For the service with no fixed cost and an increasing marginal cost, average variable cost is always below the marginal cost curve. Medicare's payment thus has a chance to be optimal for such a service given physicians with some positive weight placed on patient health benefits.

3) Medicare has to guess the total provision of the service in order to come up with an average cost estimate

Now let's examine what fixed costs mean for average-cost reimbursement. Assume MC = $50 and F = $200.



With fixed costs and constant marginal costs, average cost always exceeds marginal cost.

Average cost payment will imply powerful incentives for high quantities of care provision.

Real World MRI Example: Let's put these numbers in the context of Medicare's reimbursement for MRIs, as initially established in the late 1980s.

- Fixed Cost of MRI Scanner: $2 million (depends on quality)
- Average Variable Cost: Medicare estimates $200 (staff time administering the scan and physicians' time interpreting the image)
- Assumed utilization: 2500 (we'll talk in a moment about where that came from)

Medicare set

$$r = \frac{F}{q^{est}} + AVC$$

Medicare's estimate of utilization q^{est} was 2500 MRIs.

$$r = \frac{\$2M}{2500} + \$200$$

⇨ $r = \$800 + \$200$

⇨ r = $1000

How would you go about estimating utilization?

Medicare's estimate was based on utilization of roughly 25 hours a week. How did Medicare come up with this number and how might you want to go about it?

> **Answer:** Survey estimates of utilization.
>
> **Question:** Why would survey estimates be likely to understate subsequent utilization?
>
> **Answers:**
> 1) Subsequent utilization reflects the incentives Medicare just set up!
> 2) Because the payment rate would be smaller if the estimated quantity was large, radiologists had an incentive to understate their typical utilization patterns.
> 3) Demographics shifted in ways that increased demand for MRIs (aging of the population between the late 1980s, when the initial estimates were made, and more recent years).

In practice, utilization was underestimated by nearly 100% (50 hours/wk vs. 25 hours/wk).

Let's think about the margins:

The "correct" average cost payment would have been $400 + $200 = $600.

> Profit margin on incremental service with the "correct" payment = $400.
>
> Medicare's actual payment implied a profit margin of $800.

Three Takeaways/Things To Note:

1) Both the "actual" and "correct" average cost payment leave a significant incentive for excess utilization of existing scanners
2) The "actual" payment was an overpayment that resulted in a large incentive to invest in scanners.
3) Overpayment relative to average cost has recently been corrected through a reevaluation of the utilization assumption (this only took a couple of decades!)

Differences in MRI utilization across countries:

Exhibit 8. Supply, Use, and Price of Diagnostic Imaging in Select OECD Countries

	MRI Machines			CT Scanners	
	Devices per million pop., 2008	Exams per 100,000 pop., 2008	MRI scan and imaging fees, 2009[g]	Devices per million pop., 2008	Exams per 100,000 pop., 2008
Australia	5.6	21[d]	—[f]	56.0[b,e]	94[d]
Canada	6.7[a]	42	$824	12.7[a]	122
Denmark	—[f]	38	—[f]	21.5	84
France	—[f]	49	$436	—[f]	130
Germany	—[f]	—[f]	$839	—[f]	—[f]
Netherlands	10.4	39	$567	10.3	60
New Zealand	9.6	—[f]	—[f]	12.4	—[f]
Switzerland	—[f]	—[f]	—[f]	32.0	—[f]
United Kingdom	6.9	—[f]	$179	10.2	—[f]
United States	25.9[a]	91[a]	$1,200	34.3[a]	228[a]
Median (countries shown)	*8.3*	*40*	*$696*	*17.1*	*108*

Source: Squires (2011)

Question: What's a possible alternative to average cost reimbursement?

Answer: Use a salary payment or equipment purchase subsidy coupled with a relatively small per-service payment.

Last point: International price comparisons difficult if foreign countries don't report average cost

Supplemental Material: How do other countries finance major equipment purchases?

Table: Methods for Controlling Capital Expenditures in Hospitals

Country	System Origin	Cost Containment Method
Germany	1880s	State-set capital budgets
Netherlands	1941	Government license required for purchase
France	1945	Purchases require central approval
England	1946	Allocation from National Health Service
Canada	1946-1961	Province-set hospital budgets
Sweden	1955	County government planning
Japan	1961	Active use of fee schedule

Government plays a major role in capital-equipment acquisition in other countries.

Similar incentives could be achieved privately.

- A Health Maintenance Organization (HMO) can purchase an MRI machine and then either pay its physicians on salary or pay them through small fee-for-service payments.

- A complication: If the organization buys the machine, it needs to lock the physician into providing care exclusively for its beneficiaries. (The organization wouldn't pay for the machine and then let the physician use it on fee-for-service Medicare beneficiaries.)

Extensions: The above analysis also applies to Medicare's payments for work by different specialties.

"High-skilled" specialties require longer residencies.

Specialists are compensated for this "fixed cost" through the fee schedule. An hour of specialist time is counted as more Relative Value Units than an hour of general practitioner time.

> So the fixed cost of a longer residency is compensated through higher per-unit reimbursements, just as in the case of the MRI scanner.
>
> This highlights a conceptual overlap between Medicare's treatment of Human Capital (the investment in education) and Physical Capital (the investment in an MRI machine).
>
> Contrast U.S. financing of medical education with many foreign countries.
> - In the U.S. you pay up front and get reimbursed generously later
> - In many foreign countries medical education is highly subsidized and salaries are lower later

Further extensions:

- Medicare treats malpractice insurance premiums similarly
- Other countries front more of the fixed cost of medical education and have less generous payment afterwards