In November of 2007, *Health Affairs* (2007) selected John Wennberg as the “most influential health policy researcher of the past quarter-century.” The health care reform debate of 2009–2010 confirmed that judgment. The variations literature, in which his Dartmouth group has played so large a part, dominated discussion of policies for controlling health care costs.

In part because of that influence, discussion about cost control focused overwhelmingly on just one part of the cost equation: reducing the volume of care, rather than the price of care. As Jonathan Oberlander and I wrote during the debate, that directed attention to measures that were most likely to be unpopular, rather than measures that were more likely to be popular (Oberlander and White 2009a, 2009b). This was not only questionable politics but also badly flawed policy analysis.

In this short commentary I will begin by giving evidence about the prominence of the variations argument. Then I will give some of the reasons why I believe the following:

- The most common claims about the extent of health care system costs that are caused by unnecessary services, both in and based on the Dartmouth scholarship, appear excessive.
- The Dartmouth scholars have gone out of their way to claim their work shows that volume is the answer; implicitly and sometimes explicitly they argue that paying attention to prices is wrongheaded.
This part of their argument is not the strongest analysis within their own work and sometimes is based on only a partial view of other relevant literature.

- The emphasis on volume in the variations literature is contradicted by other evidence, which is at least as relevant to cost-control choices. The variations literature is widely interpreted as showing that the problem is paying fee-for-service. There is no ideal way to pay for medical care, but in the United States the biggest problem with fee-for-service is the specific fees per service.

- The emphasis on reducing unnecessary volume rather than lowering prices has a further flaw: we know very little about how to do the former and a lot about how to do the latter. Many of the theories generated by the Dartmouth scholars or by people interpreting their work are either not supported by evidence or face severe obstacles to practical application.

- For all these reasons, the Dartmouth scholars’ overemphasis on volume has probably done more harm than good to the cause of better controlling health care costs in the United States. This is, simply, a shame. The work makes an important contribution to our understanding of medical economics. It both suggests some of the concerns that should be targeted in price regulation and provides ideas about what to do in addition to taking the right steps on prices. The work’s misuse for policy analysis is unnecessary.

**Prominence**

The variations literature was ubiquitous in the 2009–2010 policy debate. The most widely cited advocate of the arguments was Office of Management and Budget (OMB) director Peter Orszag, who continually made claims such as, “If we can move our nation toward the proven and successful practices adopted by lower-cost areas and hospitals, some economists believe health care costs could be reduced by 30 percent—or about $700 billion a year—without compromising the quality of care” (Orszag 2009). He had long argued that the “fiscal challenge” included “the opportunity to reduce costs without impairing overall health outcomes” (Orszag and Ellis 2007).

But Orszag was preaching to the converted. When he testified to the Senate Committee on Finance on March 10, 2009, Orszag used maps from *The Dartmouth Atlas of Health Care in the United States* to demonstrate the variations and made his claims about savings; meanwhile, “around
the dais, lawmakers from both parties nodded” (Armstrong and Wayne 2009a). In June, a bipartisan group led by former Senate majority leaders Howard Baker, Tom Daschle, and Bob Dole issued a reform plan and used the variations literature to support a claim that it was possible to “create accountability for improving the overall value of health care” (Baker, Daschle, and Dole 2009).

When the House Committee on Ways and Means and the Senate Committee on Finance held hearings about reform, business leaders frequently cited the variations argument and the ubiquitous 30 percent savings statistic. Steven Wojcik of the National Business Group on Health showed the assumption that the implications of the research were known to all when he told Senate Finance, “As you know, studies at the Dartmouth Institute have shown significant variation in health care spending between regions of the United States with no difference in outcomes, only 40 percent of which can be attributed to differences in illness and price” (National Business Group on Health 2009: 6). He credited both Orszag and Dartmouth with the 30 percent savings figure, arguing, “Health care costs cannot be contained if we continue to provide and pay for large amounts of inappropriate and poor quality care” (ibid.: 1). The ERISA Industry Committee (2009) also cited the 30 percent savings claim. James A. Klein of the American Benefits Council (2009) emphasized variations but did not use the number. Ronald Williams (2009), president of Aetna, added to the chorus, testifying that “most experts agree that 30 percent of our health care is unnecessary, and yet the majority of Americans believe they don’t get the tests and treatment they need.”

The only business witness who suggested that prices might be a concern suggested action to address that concern only as a backup approach. Peter Lee of the Pacific Business Group on Health testified that the system needed to be reengineered in ways that the Dartmouth analysts and others have suggested. But he also emphasized that, compared to the government, employers were at a disadvantage in negotiating prices with providers, suggesting that Congress might eventually have to turn to “all-payer pricing or global budgeting” (Lee 2009). In contrast, Williams of Aetna objected to the idea of a public plan that would pay providers rates based on Medicare payment by arguing that “the government is missing the point—it is not how much we pay that is the problem, it is what we pay for that caused high volume consumption of health care services” (Williams 2009). Coming from Aetna, this was a self-serving and defensive argument, but it appeared to be believed by many political actors.

President Barack Obama was highly influenced by Orszag and the
variations literature (Alter 2010: 261). When he addressed cost control in public meetings, he continually phrased the effort in terms of reducing unnecessary care (Alter 2010; Nather 2009). He would insist, “That’s not rationing, it’s being sensible” (Nather 2009). Challenged as to whether doctors would be required to use “best practices . . . by law,” he dodged (Alter 2010: 261). He appears never to have understood that one of the core arguments for the so-called public plan that was pushed by liberals during the debate was to drive down medical prices. Thus in the press conference during which Obama (2010) defended his December 2010 tax-cut deal, he described the public-plan debate as being about getting “a few million people” into government coverage — so, about nothing important.

Some politicians, however, particularly conservative Democrats, did understand the price-reduction argument underpinning the public plan — and found the variations argument much more attractive. Thus the House Blue Dog Coalition declared in its “Principles for Health Care Reform” that “up to one-third of the $2.3 trillion spent on health care each year is unnecessary or duplicative” (Blue Dog Coalition 2009). Representing mainly rural areas, they saw the Dartmouth analysis as a reason to resist price regulation, for reasons discussed below.

**Excessive Claims**

There were two problems with the claims about savings that could be achieved by reducing variations. The first is that in public statements the Dartmouth scholars and others tended to exaggerate the importance of volume, even if they were more careful in their academic work (Abelson and Harris 2010). For example, Elliott Fisher testified to Congress that, “almost all of the differences in spending across both regions and academic medical centers are due to greater use of what we refer to as ‘supply-sensitive services’” (Fisher 2009a). But that’s not what the Dartmouth scholars’ own research says — it says that almost all of the variation in services is due to supply-sensitive services.

But the larger problem is the 30 percent figure itself. The Dartmouth scholars would derive it from various sources. In their very prominent 2003 articles in the *Annals of Internal Medicine* (Fisher et al. 2003a, 2003b) they used the figure but referred to a 1997 article as the source (Skinner and Fisher 1997). That article, however, only projected 20 percent savings. Perhaps they were thinking of a different article, which projected 28.9 percent savings (Wennberg, Fisher, and Skinner 2002).

The latter article, published in 2002, projected savings if costs were
lowered to the average in the lowest decile. Since the average in the lowest decile has to be at a cost somewhere below the 10th percentile, this is, in essence, assuming that over 90 percent of the regions studied have excessive costs and that the remaining small portion of the population is not unrepresentative in the least. Since it evidently is an outlier, one might at least be cautious.

The various studies also make a series of intriguing adjustments in order to distinguish volume effects from price effects. According to the notes in the 2002 study, its own price adjustments (referenced as from Dartmouth Institute 1999) were based on a surrogate price-level index rather than the actual prices paid to hospitals. This approach should not net out important factors such as disproportionate-share hospital (DSH) adjustments. The definitions of price and volume for inpatient care used in the 2008 Dartmouth Atlas do not reflect what any payer, such as the Centers for Medicare and Medicaid Services (CMS), would call the price it paid.1

A cautious person might settle for saying that costs could reasonably be reduced by reducing volume to the norm at, say, the 25th percentile. But that would not have yielded such impressive numbers. One might also look for some other assessments of the proper adjustments for illness and price. Recent research has strongly suggested that the 30 percent figure is inaccurate on these grounds as well.

These studies have not suggested an alternative figure, but they have given good reason to believe the actual dispersion of costs due to unjustified volume differences is more compact than the Dartmouth scholars report. Therefore, savings from moving higher-use regions to the low end of the distribution must be less.

Zuckerman and colleagues (2010) reduced the difference between the top and bottom quintiles for regions within states from 52 percent to 33 percent with measures of demographics, baseline health, and changes in health status. MedPAC (2009) showed that adjustments for health status, demographics, and prices (including factors such as DSH payments)

---

1. Basically the authors of the 2008 Dartmouth Atlas take hospitalizations, determine the total payment, count the number of inpatient days, divide the payment by days, and define the result as price and the number of days as volume. But Medicare doesn’t pay that way — except for outliers, it pays a fixed fee for a hospitalization. So the authors will come up with a higher price for the same actual payment to the hospital if the patient is hospitalized for three days instead of five. That can only introduce a lot of error to estimates. The scholars conclude, for example, that the price at Massachusetts General is lower than the price at Mayo St. Mary’s (Dartmouth Institute 2008a: 43–44). Given the various factors in the formulas to adjust the prospective payment system payments, this seems unlikely.
reduced dispersion to a point where only 5 percent of the units in the study (metropolitan statistical areas and rural parts of states) had costs below 85 percent of the average. In principle, health-status adjustments are likely to overadjust, because the same practice norms that lead to excessive care (in some places) could lead to overdiagnosis (Song et al. 2010). Yet even if a portion of the health-status adjustments should be discounted, the 30 percent figure seems exaggerated.

I cannot say what number would be right. Saying spending should be at around the 5th percentile but the health adjustments in the MedPAC data are correct would project a rough estimate of about 15 percent excess services. The 20th percentile would project something more like 10 percent excess. But there is surely sufficient reason to be highly skeptical of the 30 percent claim. Accepting the 30 percent claim requires one to believe that the far tail of the distribution is an appropriate level of services and that no diagnosis-based adjustment is appropriate. That is not a good basis for policy.

**Prices and the Dartmouth Research**

The Dartmouth scholars often write as if the world believes that variations in Medicare costs are due to price differentials (Dartmouth Institute 1999: 11; 2008a: 22). It would be more accurate to say that other participants in policy debate have thought controlling prices was crucial to controlling costs, rather than variations in costs.

The Dartmouth scholars should be more careful about distinguishing the two. In the 2008 *Dartmouth Atlas*, for example, they argue their analysis shows that, “most of the variation is due not to price, but to variation in the volume or intensity of care delivered” (Dartmouth Institute 2008a: 22). A few pages later they conclude that “volume, rather than price, drives reimbursement rates” (ibid.: 27). But reimbursement *rates* are not the same thing as reimbursement *variations*.

Their regressions show that volume explains much more of the variation than price does. Yet what this really shows is that Medicare has done a much better job of controlling prices than of controlling volume—a fact the Dartmouth scholars actually note (ibid.: 27). To see why this matters,

---

2. Since MedPAC did not publish the exact data, I am projecting very roughly, from eyeballing the charts. But with just over 5 percent of the population covered in regions with adjusted costs (due to volume differences only) below 85 percent of the average and about 25 percent between 85 percent and 95 percent of the average, a smooth distribution would put the break point just above 90 percent of the average.
imagine what they would have found if every price were doubled. Costs would have been twice as high, of course, but the scholars would still have found that the variation in costs was explained more by volume than prices—because the relative variation in price wouldn’t have changed. In short, to understand reimbursement rates—as opposed to variation in reimbursement rates—the acknowledgment that Medicare has done a good job of controlling prices (at least as compared to other payers) is as significant as the fact that it (like other payers) has not controlled volume well.

The Dartmouth scholars do recognize that price may play a greater role among the non-Medicare population than it plays in their studies of Medicare (see Fisher’s comments in Iglehart 2009). They should correspondingly caution political figures who insist on claiming that the findings from Medicare apply to the health care system as a whole. When they write about policy choices, the Dartmouth scholars sometimes simply ignore price. In 2009, for example, they argued that coverage for the uninsured could be financed in three ways. Two of them, however, “raising taxes” and “rationing needed care,” were not deemed “attractive.” Instead, “we should be able to reorganize and improve care to eliminate wasteful and unnecessary services” (Sutherland, Fisher, and Skinner 2009: 1227). Paying less for the services we get apparently was not an option.

Sometimes, instead, the Dartmouth scholars reject price regulation directly. In Elliott Fisher’s words, “we can’t fix the problem—and may make it worse—by focusing just on prices. Some services are clearly overpaid, but cutting prices to high-cost regions will only cause providers in those regions to deliver more services” (Fisher 2009b). Yet in most markets, as the Congressional Budget Office (CBO) explains, lower price tends to induce lower supply. The CBO reports that this appears to be the case, in fact, for Medicare home-health and skilled-nursing-facility services, and there is even some evidence that it occurs for hospital care. Physicians do appear to increase volume to offset the effects of price restraint, but the CBO estimates the volume offset as only 25 percent of the price effect (CBO 2008a; see also Medicare Actuaries 1998).

The Dartmouth analysis had one other important, though I assume unintentional, effect on cost-control deliberations in 2009–2010. The variations literature convinced rural legislators and those from other so-called low-cost regions that their providers are being underpaid. After all, their

3. That’s how regression works: it relates the range of variation in one variable to the range of variation in another.
regions are considered high value according to prestigious research! This became a major reason for opposition to a public plan, on the grounds that it would be unfair to high-value providers in some regions if the public plan paid Medicare rates—even though the price variations in Medicare are based on input costs (Wayne and Armstrong 2009; Armstrong and Wayne 2009b; Wayne and Epstein 2010).

**Evidence from Other Perspectives**

The relative importance of price and volume looks quite different if one considers investigations of cost control from other perspectives. Most obvious are the international comparisons. Research continually shows that the “variation” between costs in the United States and other countries is due far more to different prices for service than to levels of service. As Anderson and colleagues state, “It’s the prices, stupid!” (Anderson et al. 2003; Angrisano et al. 2007; Ginsburg 2008).

Within the United States itself, Medicare has controlled costs better than private insurers have (MedPAC 2008), again, because of its better control of prices. In fact, the same reforms that helped Medicare improve its cost control also reduced overall Medicare cost variation by reducing variation due to price, leading to the Dartmouth findings. Medicare payment reforms that moved from cost- or charge-based systems to the prospective payment system and the resource-based relative value scale substantially reduced variation within Medicare hospital and physician payments: “Under cost- and charge-based systems, providers had considerable influence over payment rates, which in turn allowed for substantial geographic variation in payments” (CBO 2008a: 5).

Analysis of cost differentials within U.S. private insurance markets also shows that price trends are extremely important. Conventional wisdom holds that something called “managed care” controlled costs by reducing volume in the mid-1990s, and costs soared when volume controls were loosened. But the evidence from the Center for Studying Health System Change as well as other studies is quite different: both the period of cost control and the collapse of cost control were much more strongly related to payers’ ability or inability to restrain prices (White 2007). In the current market, provider “pricing power” drives up costs, according both to insurers and to neutral analysts (CBO 2008a; Ginsburg 2010). If they look at prices closely, strong advocates of delivery reform recognize that prices

---

4. For a more thorough review of the case for price regulation, see White 2009.
are a major policy issue and, perhaps, a lever. In Bob Berenson’s words (Iglehart 2009), “To create the conditions to move away from fee-for-service, you need to pay attention to the distortions in the current Medicare fee schedule, which then get compounded even more in private fee schedules.” He adds that “we’ve got the prices really wrong” (ibid.).

**Lack of Effective Responses**

None of these criticisms would be all that important if there were obvious and acceptable policy choices to reduce variation in services. Yet there are good reasons why the CBO (2008b) projected hardly any savings from delivery reforms that were on the table during the 2009 debate. As Katherine Baicker has said, it sounds easy “if we can say, ‘here’s the health care that does a lot of good, and here’s the health care that’s kind of wasted—let’s not do that.’ . . . But of course it’s not so straightforward to say which use of health care resources is really driving improvement in health and which resources aren’t doing such a good job . . . It would be pretty hard to write down a set of public policies that pushes resources towards this while withholding resources from that . . . Even providers don’t know that. There’s a huge gray area of medicine” (Epstein 2009: 6)

This lack of knowledge is one of the Dartmouth group’s core and accurate points. But their answers have been wishful at best. In one argument they call for a program in which organized delivery systems voluntarily contract with CMS to pursue so-called shared savings in Medicare while research to color the gray areas is conducted. Then “after several years” there would be a transition “to prospective payment based on successful completion of the research agenda” (Dartmouth Institute 2008a: 106). Unless “several” means something like “thirty,” this is a highly optimistic view of the research enterprise.

I understand the attraction of a world of competing “organized delivery systems” (Dartmouth Institute 2008b: i)—the goal that has attracted many health policy specialists since the 1980s. It did not happen in the 1990s because it was a poor business model in many ways (White 2007). These policy specialists do not want to leave transformation to the market; instead they want to use Medicare contracting to encourage or force systems to integrate and then downsize capacity, of both equipment and personnel (Dartmouth Institute 2008b: 5). While this approach would likely save money — so long as providers in short supply did not exploit the situation to demand higher prices, of course — the political objections could make price regulation seem easy in comparison.
In fact, one may find evidence in the Dartmouth reports’ own findings to question whether “organization” and reducing “fee-for-service incentives” can control costs. “Organized practices such as the Mayo Clinic, Billings Clinic, and Kaiser Permanente,” the Dartmouth scholars argue, “would be well positioned to respond to the incentives of a shared savings program, by expanding within markets where they already have a presence, and entering new markets through the purchase of hospitals and recruitment of resident physicians” (Dartmouth Institute 2008b: 6). Yet they show that the Mayo Clinic has not, in fact, brought a consistent approach into markets as they have expanded: “We find no evidence that providers in these systems use a distinctly Mayo Clinic strategy for allocating resources and managing chronic illness” (Dartmouth Institute 2008a: 61).

Another supposed example of savings from having an “organized delivery system” is the Cleveland Clinic, which is highlighted by comparison to other “top-five” hospitals in the 2008 Dartmouth Atlas. It was widely argued that the Clinic had lower costs than, say, Massachusetts General Hospital because its physicians, being salaried, had less incentive to proliferate services (Adler 2009). Yet the right comparison to the Cleveland Clinic, in order to control for nonorganizational factors, would be as similar as possible. A mile down Euclid Avenue from the Clinic is another huge and prestigious academic medical center, University Hospitals of Cleveland (UH). In the study, UH’s costs per patient ($55,643) were nearly identical to the Cleveland Clinic’s ($55,333). Moreover, the clinic had a Dartmouth hospital-care intensity index ranking in the 50th percentile, while UH’s was around the 25th (Dartmouth Institute 2008a: 109–111). Yet UH physicians were typical academic medical center doctors. In short, it wasn’t salary; it was Cleveland.

The Veterans Health Administration (VHA) is another highly organized delivery system; it does not receive fees per service, yet it also exhibits “substantial variation in patterns of clinical practice despite the fact that [the VHA’s] management tracks providers’ compliance with medical guidelines for many medical conditions . . . The implication is that local norms can influence practice patterns, even in a relatively centralized system that places a strong institutional emphasis on adherence to clinical guidelines for care” (CBO 2008a: 10).

The variations literature identifies striking features of the medical world that have important implications for policy. But it certainly does not offer compelling ideas about what to do about the variations. Politicians do no better, of course. President Obama emphasized how research would make decisions clearer, even though, as any good reporter could figure
out, this was quite unlikely within a plausible budget horizon (Nather 2009). Orszag claimed that health information technology, research, wellness, and “changes in financial incentives for providers so that they are incentivized rather than penalized for delivering high-quality care” would do the trick, simply assuming away difficulties such as the weakness of quality measures (Orszag 2009).

The variations argument does, however, direct politicians toward making the kind of claims that are likely to scare voters. As Mark McClellan pointed out, “It becomes a real challenge to on one hand explain how peoples’ relationships with their doctors won’t change and they get to keep their coverage secure, while at the same time stressing how important it is for the system to change in order to stop this unsustainable trajectory in spending” (Bettelheim 2009).

Conclusion

The Dartmouth enterprise has made numerous important contributions to health policy discourse.

The first is political. The variations work clearly documents the degree to which medical practice is not based on rationalized professional consensus. The Dartmouth project in this sense is part of an international enterprise that justifies having a profession of health services research. This has been central to challenging medical clinicians’ sway over much of health care policy— not just in the United States but abroad (Bevan et al. 2004). From a narrow professional perspective alone, this might justify *Health Affairs’* award to Professor Wennberg.

Second, the variations literature demolishes a dominant and fatalist argument among economists about the growth of health care costs. As the Dartmouth scholars put it, the findings “call into question the notion that additional growth in health care spending is primarily driven by advances in science and technology and that spending more will inevitably result in improved quality of care” (Fisher et al. 2003a). Clearly, technology does not implement itself; a great deal of human choice is involved— which means that policy could also matter.

Third, the somewhat idealized distinctions between supply-sensitive care, preference-sensitive care, and care that is well justified by an evidence base provide a fine way to think about research needs; patient self-determination; and, in general, what policies might be applied for what choices.

Fourth, the emphasis on supply ought to bring back into U.S. debate the
question of capacity regulation—which is taken for granted as an important part of cost control in most other countries.

Fifth, the work does show that more research about practice could be a good idea—so long as expectations about what that research could accomplish, and when, are more reasonable. It adds to other lines of research that suggest quality should and could be improved.

Yet the overemphasis on the variations finding during the recent reform debate should be reversed, and it would be very helpful if the Dartmouth scholars would agree with that.

Volume regulation alone cannot control costs. In fact, the evidence from insurance markets is clear: the capacity reductions that the Dartmouth scholars advocate will only increase providers’ market power and thus drive up prices, without strong price regulation. It is not at all clear that higher prices for fewer services would be either less expensive or better value for money.

The importance of limiting prices should not be a matter of dispute. At the same time, it is clear that policy needs to find the right bundles for payment. There are risks in all choices, but there is no particular reason why scholars who focus on volume and those who focus on price should assess the choices differently.

Another logical approach to address both prices and volume is volume-adjusted prices: if the quantity of a specific service goes up faster than quantities of other services, the former should face more price constraint. In general, greater sales should mean a lower price per unit, and lower prices would capture efficiencies of production (as with advances in eye surgery). Lower prices would also discourage excessive investment in equipment (Ginsburg 2008). The key is to take this approach at the level of specific services—unlike the across-the-board adjustments in the Medicare sustainable growth rate mechanism. As Bob Berenson and many others have argued, the relative values in the Medicare schedule, which is largely copied by private payers, badly need adjustment (Iglehart 2009). More regular volume adjustment would be part of such a practice.

In addition, the Dartmouth scholars claim, as do many others, that there is an oversupply of specialists and undersupply of generalists. One obvious way to address this imbalance would be to adjust the fees, paying specialists less for their services and generalists more—though not necessarily as much more as specialists are paid less!

The Dartmouth analysts may not be interested in another approach, namely, having lower prices for services in areas with higher levels of service. They appear to prefer having some form of budgetary pressure in
high-cost areas. But the effect would be the same in the long run: incomes go down, some capacity gets squeezed out. If Medicare lowered its fees in particularly high-cost areas, that would not be ideal and fair to all providers, but it would surely reduce costs and provide incentives to reduce supply—and could be implemented much more quickly than any of the Dartmouth ideas.

In fact, getting the fees right is the place to start in providing incentives for providers to decide they’d be better off trying to increase the value of care in an integrated way. If the Dartmouth group wants to reduce variations and control costs, they should recognize that better price regulation is a step toward what they want as well.

References


Fisher, E. S. 2009a. The Implications of Regional and Provider-Specific Variations in Medicare Spending for Medicare Payment Reform. Testimony to Committee on Ways and Means, U.S. House of Representatives, April 1.


