

CHAPTER 14

Unemployment Insurance, Disability Insurance, and Workers' Compensation



James Andrews/Stock Getty Images

14.1 Institutional Features of Unemployment Insurance, Disability Insurance, and Workers' Compensation

14.2 Consumption-Smoothing Benefits of Social Insurance Programs

14.3 Moral Hazard Effects of Social Insurance Programs

14.4 The Costs and Benefits of Social Insurance to Firms

14.5 Implications for Program Reform

14.6 Conclusion

APPENDIX TO CHAPTER 14

Advanced Quasi-Experimental Analysis

Questions to keep in mind

- How do the unemployment insurance, disability insurance, and workers' compensation programs fit within our general social insurance framework?
 - What are the consumption-smoothing benefits and moral hazard costs of these programs?
 - How does the evidence on behavioral responses suggest that these programs be reformed?
-

The Covid-19 pandemic created a public health catastrophe in the United States, with more than 40 million cases and 650,000 deaths as of August, 2021.¹ It also

created an economic catastrophe and, in doing so, placed enormous stress on our nation's Unemployment Insurance (UI) system—administratively, fiscally, and politically. The number of new weekly unemployment claims peaked at more than 6 million in early April 2020, more than six times the peak of the new weekly claims during the 2007–2009 Great Recession.

As described in more detail in this chapter, the UI system levies payroll taxes on employers to finance cash benefits to those who are laid off from their job. Benefits are available only to workers who have worked substantial amounts in the previous year in “covered” employment, which includes most traditional jobs. When those workers lose their jobs, they are entitled to a benefit which is a share of their pre-unemployment earnings, up to some maximum amount. These benefits typically last for 26 weeks, although they have been extended for longer periods in previous recessions (most notably up to 99 weeks during the Great Recession).

The Coronavirus Aid, Relief, and Economic Security (CARES) Act passed in the early months of the pandemic made three significant changes to UI, all of which were 100% federally financed, unlike the traditional employer payroll-based financing of UI. The Pandemic Emergency Unemployment Compensation (PEUC) authorized up to 13 additional weeks of federally financed UI benefits for those who exhaust their UI entitlement, mirroring changes made during previous recessions. The Pandemic Unemployment Assistance (PUA) program dramatically extended the availability of UI to individuals who have been traditionally excluded from the UI system, such as the self-employed, independent contractors, and gig economy workers, for 39 weeks. And the Federal Pandemic Unemployment Compensation (FPUC) program provided an additional \$600 benefit per week for all recipients of UI, through July 31.²

Almost as soon as these benefits were passed, they became a source of controversy—particularly the FPUC program. Supporters maintained that the FPUC was a necessary lifeline that was preventing catastrophe for millions of unemployed workers in a shut-down economy. Senate minority leader Chuck Schumer argued that without the policy, “millions of American families [would] have their legs cut out from underneath them at the worst possible time.”³ Indeed, combined with the \$1,200 Economic Impact Payment stimulus checks that were sent out to many Americans as part of the CARES Act, these UI extensions led to an actual *reduction* in poverty in the United States during this period. [Han et al. \(2020\)](#) found the poverty rate fell by 2.3% from February to April, with government programs—including extended UI benefits and the Economic Impact Payments—responsible for the full

decline in poverty. This decrease in poverty occurred despite a 14% reduction in employment in April—the worst month on record.

But opponents argued that tying this substantial benefit to being out of work would remove any incentive for the unemployed to search out new jobs. Indeed, the \$600 bonus was quite substantial compared to previous earnings for many workers, and when combined with existing UI benefits resulted in 76% of workers actually making *more* while unemployed than they did while at work.⁴ As a result, the relief bill's UI benefits faced opposition in Congress. Senator Rick Scott stated “we cannot create an incentive not to work. We cannot be paying people more money on unemployment than they get paid in their job.”⁵

This argument came to head during the summer of 2020, as the FPUC program neared exhaustion. The Democrat-controlled House of Representatives passed the HEROES Act, which would extend the \$600 bonus for six more months, through early 2021. In response, the Republican-controlled Senate developed the HEALS Act, which would have instead provided only \$200 per week through September, followed by a payment replacing 70% of lost wages through the end of the year.⁶ These differences of opinion could not be resolved, and the FPUC ran out at the end of July. Both sides continued to spend the rest of that (election) year fighting over extending benefits, until the \$900 billion relief bill signed on December 27, the day after the PUA and PEUC programs expired, provided a \$300 weekly supplement to benefits for jobless workers through mid-March, 2021.⁷ President Biden has proposed raising that supplement to \$400 and extending it through March 2021; that benefit was then extended by Congress through September 2021 before ending.

This debate illustrates that public provision of insurance for unemployment is considered an important role for government in the United States. [Chapter 12](#), however, taught us that public finance economists must step back and ask why there is a role for the government in providing this insurance. We must also consider whether the benefits of extending UI (the additional consumption smoothing for displaced workers) exceed the costs (the moral hazard effects on unemployment).

In this chapter, we discuss in detail the issues that arise in providing UI and two other major forms of social insurance: workers' compensation insurance against on-the-job injury and disability insurance (DI) against career-ending disabilities. These are three of the largest social insurance programs in the United States, and they share many common features. Benefits are conditioned on the occurrence of some adverse event (unemployment, disability, injury) and are based on past earnings on

the job. Most important, benefits are based at least partly on unobservable or difficult-to-verify conditions, such as job search effort and the true inability to work due to disability or injury. The unobservability of some aspects of the adverse event raises moral hazard problems that must be considered in program design.

We begin the chapter by describing the structure of these three programs and comparing their similarities and differences. The next section presents the empirical evidence on the costs and benefits of these programs and uses the framework from [Chapter 12](#) to discuss the implications for optimal program benefit levels. We then look at the impact of these programs on employer decision making over issues such as worker layoffs and workplace safety. We conclude with a discussion of the implications of these findings for program reform.

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
reibaker@live.unc.edu

In this chapter, we discuss in detail the issues that arise in providing UI and two other major forms of social insurance: workers' compensation insurance against on-the-job injury and disability insurance (DI) against career-ending disabilities. These are three of the largest social insurance programs in the United States, and they share many common features. Benefits are conditioned on the occurrence of some adverse event (unemployment, disability, injury) and are based on past earnings on the job. Most important, benefits are based at least partly on unobservable or difficult-to-verify conditions, such as job search effort and the true inability to work due to disability or injury. The unobservability of some aspects of the adverse event raises moral hazard problems that must be considered in program design. We begin the chapter by describing the structure of these three programs and comparing their similarities and differences. The next section presents the empirical evidence on the costs and benefits of these programs and uses the framework from [Chapter 12](#) to discuss the implications for optimal program benefit levels. We then look at the impact of these programs on employer decision making over issues such as worker layoffs and workplace safety. We conclude with a discussion of the implications of these findings for program reform.

14.1 Institutional Features of Unemployment Insurance, Disability Insurance, and Workers' Compensation

As with Social Security, it is difficult to discuss the economic implications of these social insurance programs without first understanding how they work. In this section, we provide a brief overview of the institutional features of these three programs.

Institutional Features of Unemployment Insurance

Unemployment insurance (UI) is a federally mandated program that is administered by each of the 50 U.S. states and the District of Columbia, which are free to set their benefit levels and other aspects of the program.⁸ That UI is administered by the states is an excellent feature of this program for the purposes of empirical economics because economists can learn about the consumption-smoothing and moral hazard effects of UI by studying how these effects vary as the generosity of state benefits vary. Indeed, the large empirical literature discussed in this chapter derives its success from the wide variation in UI programs across the states.

unemployment insurance (UI)

A federally mandated, state-run program in which payroll taxes are used to pay benefits to workers laid off by companies.

Unemployment insurance is financed by a payroll tax that is levied on employers and averages 0.51% across the states.⁹ As with the Social Security payroll tax, this tax is paid only on earnings up to a certain level, a fairly low level in most states. The UI payroll tax is partially experience-rated; that is, it rises as firms have more layoffs, but not on a one-for-one basis. Thus, firms with twice as many layoffs do not typically pay twice as much in payroll taxes.

partially experience-rated

The tax that finances the UI program rises as firms have more layoffs, but not on a one-for-one basis.

An important feature of UI is that not all those out of work qualify for benefits. To be eligible, workers must meet four criteria. First, workers must be employed in covered employment, which is work done by an individual for an employer rather

than as a sole proprietor; this covers 97% of non-farm workers in the U.S. economy.¹⁰ Second, workers must have earned a minimum amount over the previous year; in most cases, states require minimum earnings in several previous calendar quarters before individuals are eligible for benefits. Third, workers cannot quit or be fired for cause; UI is available only to those who are laid off from their jobs for economic reasons. Finally, workers have to be actively looking for work and willing to accept a job comparable to the one they lost. Even if eligible, individuals don't receive benefits automatically. They must go to the UI office and enroll in the program, and they must show evidence that they are looking for a new job.

Between eligibility requirements and the effort required to claim UI, only 28% of unemployed workers actually received UI benefits in 2019.¹¹ Even among those eligible, receipt of benefits is only about three-quarters.¹² The reasons for this lack of participation among those eligible for benefits are unclear, but they are usually divided into two issues. The first is information: some individuals may not understand that they are entitled to UI when they lose their jobs. The second is the stigma attached to being unemployed: individuals may be reluctant to sign up for what they see as a government handout, particularly because they would actually have to go to a public office and state that they can't find a job.

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
reibaker@live.unc.edu

The final consideration has to do with targeting the benefits to those who need them the most. Workers away from work (due to either unemployment or injury) for the longest period may be most in need of assistance from the government in finding a new job; if they could readily find one, they would have done so already. Under this logic, cutting benefits off after some fixed time is exactly the wrong thing to do. Given these three considerations, the optimal duration of social insurance benefits is difficult to compute, which may explain the differences in structures across similar industrialized nations.

[Kolsrud et al. \(2018\)](#) incorporate these effects into the context of UI duration in Sweden. They use the fact that the replacement rate available to individuals in Sweden varies with the length of the unemployment spell to show how both consumption smoothing and moral hazard vary with spell duration. In fact they find that moral hazard is larger for benefits paid earlier in the spell and that consumption smoothing grows as the spell lengthens. This finding strongly suggests that the optimal time pattern of UI benefits would not be falling, and may even be rising, over time. ■

14.2 Consumption-Smoothing Benefits of Social Insurance Programs

As discussed in [Chapter 12](#), one determinant of the optimal size of social insurance programs is the extent to which they crowd out self-insurance rather than providing consumption smoothing. Unfortunately, there is relatively little evidence on the consumption-smoothing implications of these programs. The most direct study is that of [Gruber \(1997\)](#), who measured the consumption-smoothing benefits of UI. He found that individuals are not fully insured by other sources against the income loss of unemployment, that their consumption falls significantly when they lose their jobs, and that higher levels of UI do lessen the negative effects of this fall. Gruber's study also found, however, that the role of UI in consumption smoothing is limited: each dollar of UI benefits only reduces the fall in consumption by about \$0.30. UI smooths consumption in a limited way because other forms of self-insurance enable unemployed workers to maintain part of their pre-unemployment consumption.

Another way to phrase this conclusion is that UI *crowds out* other forms of income support when people are unemployed. Other research shows directly that crowd-out occurs: when individuals face more generous UI benefits, they save less, and their spouses are less likely to work when they lose their jobs. These studies estimate that crowd-out of UI is about \$0.70 on the dollar, consistent with the consumption findings just noted. That is, these findings imply that, of each dollar of UI, \$0.30 goes to increase consumption, and \$0.70 crowds out other forms of insurance that individuals were using for their unemployment spells.²⁴

Of course, the extent of consumption smoothing varies across individuals based on the extent to which they can self-insure. [Rothstein and Valla \(2017\)](#) studied those who had run out of extended unemployment insurance, so that they had been unemployed for a long time and likely had depleted their self-insurance. For this population, they found much more consumption smoothing, with a large rise in poverty and even an increase in self-reported disability following exhaustion of benefits. [And Ganong and Noel \(2019\)](#) show that spending falls dramatically even after predicted exhaustion of UI benefits, suggesting that the consumption-smoothing gains from extending UI benefits are four times larger than from raising UI benefit levels.

Evidence of this type for disability insurance is provided by [Autor et al. \(2015\)](#). They examined the case of Norway and looked at the impacts on consumption of being

denied disability insurance. They found that about 60% of the reduced income from being denied DI was translated to lower consumption; that is, about 40% of DI benefits were crowded out. The researchers also found that most of the crowd-out occurred through increased spousal labor supply when DI was not available. These findings indicate that there is some crowd-out for DI but that it is much less than for UI. This is consistent with evidence in [Meyer and Mok \(2018\)](#), who showed that those who become disabled see a sizeable drop in consumption, and [Deshpande et al. \(2019\)](#), who find that those who receive DI see the odds of bankruptcy fall by 30% relative to those who are rejected.

There is no parallel evidence on the consumption-smoothing properties of WC, and it is unclear what to expect for this program. On the one hand, true on-the-job injuries are probably even more unexpected than is unemployment and can often last longer. On the other hand, many claimed on-the-job injuries are not true injuries and therefore not a “surprise,” according to the moral hazard evidence presented next, so individuals may be well prepared to smooth their consumption.

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
reibaker@live.unc.edu

14.3 Moral Hazard Effects of Social Insurance Programs

In contrast to the small amount of research that has been done on consumption-smoothing effects of these social insurance programs, there is a large literature on their moral hazard effects. In this section, we review that evidence.

Moral Hazard Effects of Unemployment Insurance

The major moral hazard effect of UI is seen in the duration of unemployment spells. The question of how UI affects the odds of becoming unemployed is a separate one that we take up when discussing the behavior of firms later in this chapter.

There is a voluminous economics literature on the effect of UI benefit levels on unemployment durations. The most convincing studies in this literature use the quasi-experimental approach discussed in [Chapter 3](#). In particular, because UI programs differ so much across states, researchers can use differences in benefits across state UI systems to measure the response of unemployment durations to benefit levels. Such analyses have suggested that higher benefits lead to a large increase in unemployment durations; the best estimate, reviewed in the Empirical Evidence box, suggests that each 10% rise in the amount of UI benefits leads to unemployment durations that are 8% longer.

EMPIRICAL EVIDENCE

Moral Hazard Effects of Unemployment Insurance

Perhaps the best illustration of the quasi-experimental approach to studying the moral hazard effects of UI is [Bruce Meyer's \(1989\)](#) classic study ([Figure 14-3](#)). Suppose the UI program in New Jersey pays a maximum benefit of \$350 per week, and that any worker with earnings above \$700 per week is eligible for the maximum benefit. New Jersey then changes its UI program so that the maximum benefit is \$400 per week. After the change, any worker with earnings above \$800 per week (workers in group H) is eligible for the new maximum benefit. Workers with earnings above \$800 per week have therefore seen an increase in their benefits of \$50 per week.

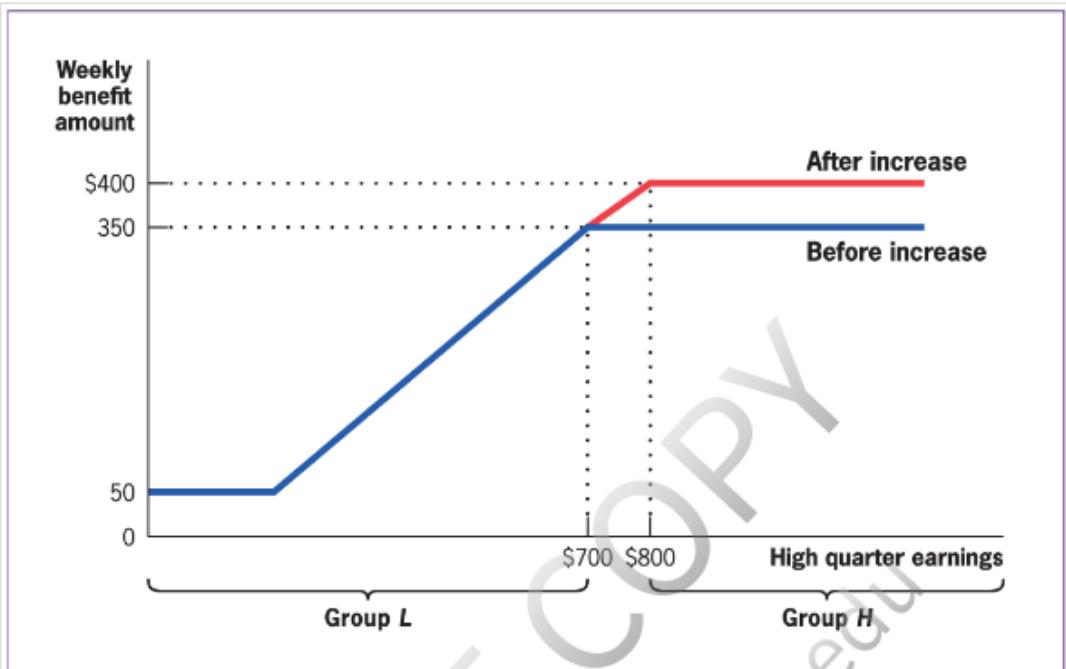


FIGURE 14-3 Quasi-Experiments for Studying the Effect of UI Benefits on Unemployment Durations •
The treatment state of New Jersey originally provides a maximum benefit of \$350 for workers earning more than \$700 per week. It then raises its maximum benefit to \$400, which applies to all workers earning more than \$800 per week (group H). Workers earning less than \$700 per week (group L) are not affected by this policy change.

Data from: Adapted from [Meyer \(1989\)](#), Table 1.



This policy change sets up a natural quasi-experiment. Workers in group H in New Jersey (the treatment group) can be compared to workers with the same earnings (\$800 or more per week) in another state that did not increase benefits, such as Pennsylvania (the control group). By doing this comparison, we can control for any changes over time that might affect unemployment durations, for example, a recession that affects both New Jersey and Pennsylvania. The change in unemployment duration in the control state (Pennsylvania) is a measure of how much the change in duration is due to the recession. The difference between the treatment state and the control state accounts for the recession effect and provides a causal estimate of the impact of the benefit change on the duration of unemployment spells.

To accomplish this analysis, Meyer computed the difference-in-difference estimate shown here:

$$\begin{aligned} \text{Duration (treatment, after)} - \text{Duration (treatment, before)} &= \\ &\quad \text{Treatment effect} + \text{Recession effect} \\ \text{Duration (control, after)} - \text{Duration (control, before)} &= \\ &\quad \text{Recession effect} \\ \hline \text{Difference} &= \text{Treatment effect} \end{aligned}$$

By comparing the change in duration in the treatment state to that in the control state, Meyer rid the estimates of the impact of changes over time, such as a recession. In doing so, he found that a 10% rise in benefits led to an 8% increase in the length of durations.

As Meyer noted, however, this difference-in-difference approach may not provide a truly causal estimate. Suppose that the recession has a different effect in different states; in particular, New Jersey is hit hardest, which may even be the reason that it raised its benefits. In this case, using Pennsylvania as a control would not be sufficient to eliminate any bias because unemployment durations in New Jersey would have risen more due to the steeper recession in New Jersey, not due to higher benefits. If the treatment and control states are affected differently by the recession, bias would continue to be a problem in this cross-state comparison.

To try to eliminate the remaining bias in this comparison, Meyer suggested an additional control: group *L* in [Figure 14-3](#), low earners in New Jersey, for whom benefits did not change. Because this group is in the same state as group *H*, they are subject to the same (New Jersey-specific) recession and, thus, to the same source of potential bias from that recession. So, we can compute a similar difference-in-difference estimate to Meyer's original, but using workers in group *L* in New Jersey as the control, rather than using group *H* in Pennsylvania as the control. Meyer undertook both these checks on his results (comparing group *H* to both group *L* within the same state and to group *H* in other states), and his conclusions were unchanged. (Note that it is possible to use both checks simultaneously in an effort to rid the estimates of bias; this approach is described in detail in the appendix to this chapter.)

There is also some randomized trial evidence to draw on in evaluating the impact of UI benefits on unemployment durations. In the 1980s, a number of states pursued "reemployment bonus experiments," in which a treatment group of individuals on UI was offered a bonus if they found a job more quickly and in which a control group of UI recipients was not offered this bonus. For example, in the Illinois experiment, treatment UI recipients were offered cash bonuses of \$500 (equivalent to 4 weeks of UI benefits for the typical worker at that time) if they found a job within 11 weeks of filing for UI. These experiments generally showed a significant decline in unemployment durations of treatments, relative to controls, when the bonuses were offered; on average, across experiments, the offer of such bonuses caused unemployment durations to decline by about half a week, or 3% of average duration.²⁵ Thus, we can confirm from a randomized trial that unemployment durations do indeed respond to benefit levels.

Normative Implications of Longer Durations

The clear conclusion from empirical research is that higher UI benefits have a significant causal effect on unemployment durations. The normative implications of this positive finding are unclear, however, and depend on what the worker is doing during a longer unemployment duration. Suppose that unemployed workers can find jobs relatively easily but that when benefits are high enough, they delay taking those jobs while they take extra weeks of leisure. This outcome would imply that the increase in duration is inefficient—both because workers are not contributing to the labor market when choosing leisure over work and because higher taxes must be levied on productive workers to finance these benefits.

If workers spend this time looking for better jobs, however, longer durations might be an efficient outcome of UI. The productivity of any worker in any job will depend on how well the worker is suited to that particular job. Workers who are highly

productive in some jobs may be unproductive in others. Moreover, when workers with specialized skills lose their jobs, it may take them some time to find new ones. Taking the first job that they are offered could be inefficient. Society's productivity will clearly not be maximized if brain surgeons end up working in the fast-food sector! If UI is increasing duration by subsidizing effective (but time-intensive) job searches, then society may gain from the improved **job match quality**.²⁶

job match quality

The marginal product associated with the match of a particular worker with a particular job.

How can we distinguish whether UI subsidizes unproductive leisure or productive job searches? The best way to do so is to study the quality of the post-UI job matches. If higher UI benefits are leading to longer durations because of more productive job searches, then we should see higher benefits leading to better job matches. But if higher UI benefits are simply subsidizing leisure, then there will be no better job matches when benefits increase because individuals will just take jobs they would have taken without the benefits increase (with the benefits increase simply raising the amount of leisure before that job is taken).



© Frank Cartier/CartoonStock.com

"I do intend to seek employment, but it will be at a time and place of my own choosing."

"I do intend to seek employment, but it will be at a time and place of my own choosing."

While it is difficult to measure the quality of job matches, jobs' wage rates should be some indication of their quality. A finding that higher UI leads to higher post-unemployment wages would indicate that job matches are better. In fact, however, a series of international studies have shown that this is not the case.²⁷ Thus, it appears that higher UI benefits are not leading to better job matches and that UI has a significant moral hazard cost in terms of subsidizing unproductive leisure.

Evidence for Moral Hazard in DI

There have been a number of studies on the degree of moral hazard in the DI program. While these studies take different approaches, as reviewed below, the

Program. While these studies take different approaches, as reviewed below, the

common thread is assessing whether a more generous DI program, relative to other labor market opportunities, leads to more use of the program and/or less labor force participation. If there is no moral hazard, and individuals using these programs have disabilities which truly prevent them from working, then use of the programs (and labor supply) should be unaffected by program generosity; individuals are either too severely disabled to work or not. If program generosity has a significant effect on the incidence of disability and the level of work effort, then these programs may have moral hazard effects.

The literature on moral hazard in DI grows out of the facts shown in [Figure 14-4](#) from [Parsons \(1984\)](#). From 1957—when the DI program was introduced—until 1980, there was a tremendous concurrent expansion in the number of males age 45 to 54 receiving DI benefits and the number of men in that age group dropping out of the labor force. This striking correspondence suggests that DI plays a large role in reducing labor force participation, which is consistent with moral hazard. It is hard to draw strong conclusions from this evidence, however, because many other things were changing in the 1960s and 1970s that may have led older men to work less. For example, many of the disabled could have been World War II veterans who, because of their war injuries, may have been less physically able to work at older ages and so would have dropped out of the labor force even if there were no DI program. In addition, the growth of private pension programs over this period may have made retirement more attractive.

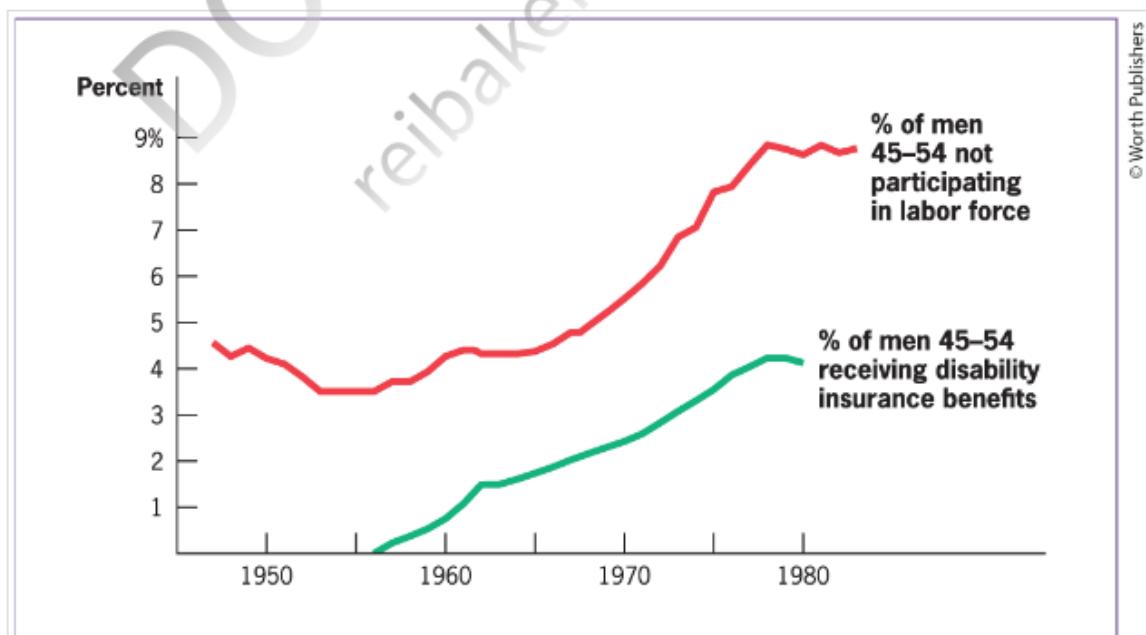


FIGURE 14-4 Labor Force Nonparticipation of Older Men and Growth in DI • There is a striking correspondence between the growth in the DI program from the mid-1950s to the mid-1970s and the rise in nonparticipation rates of men age 45 to 54 during this same period. This correspondence may indicate that the

nonparticipation rates of men age 45 to 54 during this same period. This correspondence may indicate that the

availability of DI induced older men to leave the labor force, but other factors may also explain the correspondence.

Data from: [Parsons \(1984\)](#), Figure 1.



A second source of evidence for a moral hazard effect of DI comes from the role of the business cycle. A number of studies have documented that disability applications rise sharply during recessions; one recent study found that the Great Recession led to a 6.7% rise in DI applications.²⁸ It seems highly unlikely that more individuals become disabled in recessions than in other times, so this finding is consistent with a moral hazard effect: workers are out of work anyway due to the recession, so the five-month waiting period for DI is less of a barrier, and they are willing to take a chance at applying even if they are not truly disabled.

A third source of evidence for moral hazard in DI is that individual labor supply and DI application decisions depend on the stringency of screening in the DI program. If individuals applied only when truly disabled, then the stringency of screening should affect the rate of DI acceptance, but not the decision to leave the labor force and apply. Yet a number of U.S. studies have found that both rates of application ([Parsons, 1991](#)) and rates of labor force exit ([Gruber and Kubik, 1997](#)) rise when screening is more lenient. A study of a large-scale experiment in the Netherlands, where screening intensity was increased in 2 regions and not in 24 others, found that more intense screening reduced both long-term absenteeism and DI applications ([van der Klaauw, 2006](#)). Recent studies, reviewed in the Empirical Evidence box, show that there is very substantial work capacity remaining in those who are denied disability insurance.

EMPIRICAL EVIDENCE

Disability Insurance Screening and Labor Supply

As we've seen, the most difficult issue facing disability insurance systems is determining who is truly disabled. Countries take different approaches to answering this question. But recent evidence from multiple nations reveals a consistent conclusion; those who are denied DI are often capable of gainful employment.

Assessing the work capacity of those denied DI is very difficult because, as long as screening is somewhat effective, those denied DI will be in better health than those who remain on the program. Thus, there is a clear bias in comparing those who receive DI and those who are denied. Two recent studies, however, have posed clever quasi-experimental methods to overcome this limitation.

For the United States, evidence comes from [French and Song \(2014\)](#). When individuals appeal their initial DI decisions, they go before an administrative law judge (ALJ) for a redetermination of their initial case. French and Song noted that there is substantial variation in the rate at which ALJs do and do not approve appeals: being assigned to a judge who is one standard deviation more lenient leads to a 15.3% rise in the odds of the application being allowed.²⁹ Most importantly, assignment to judges is completely random and unrelated to the severity of the injury or other characteristics of the applicant. Therefore, individuals who are assigned to “tough” judges are randomly more likely to end up without DI. These variations in the DI system allowed French and Song to ask: If an applicant is denied DI, due to random assignment to a tough judge, are they more likely to end up working? If those who apply for DI are truly unable to work, then there should be no effect on work; those denied should be no more able to work than those who are approved to receive the DI.

French and Song found that being assigned to a tougher judge led to a higher rate of denial—and a higher rate of working. Those who were randomly denied were 26% points more likely to be working compared with those who were approved. This finding suggests that a substantial minority of those receiving DI could be working if denied the insurance benefit.

[Borghans et al. \(2014\)](#) used an alternative approach to study a major DI reform in the Netherlands. The Netherlands has a more generous DI program than the United States, with more lenient standards of receipt (i.e., the applicant does not have to be as significantly disabled to be approved for benefits) and higher benefits. In addition, benefits vary by measured level of disability, with the most disabled getting higher benefits. In 1993, the system was reformed to reclassify disabilities as less severe than previously classified; some individuals with severe disabilities were reclassified as less severe (and therefore received lower benefits), while others with less severe disabilities were reclassified as not disabled. All individuals on DI were to be reexamined as part of the reform, but these new tougher standards would apply only to those age 45 and younger. This sets up a “regression discontinuity” design of the type discussed in [Chapter 6](#): we can examine whether there is a noticeable change in behavior exactly at age 45.

The results of this analysis are illustrated in [Figures 14-5](#) and [14-6](#), first for DI benefit amounts and then for earnings. Each point in the figure represents the average benefits/earnings for a birth cohort at a given age (e.g., for those born in 1944 at age 45). The x axis in each figure shows ages, while the y axis shows either DI benefit amounts ([Figure 14-5](#)) or earnings ([Figure 14-6](#)), in euros. The solid lines on either side of age 45 show a regression prediction of the relationship between age and benefits/earnings separately for those below and above that age.

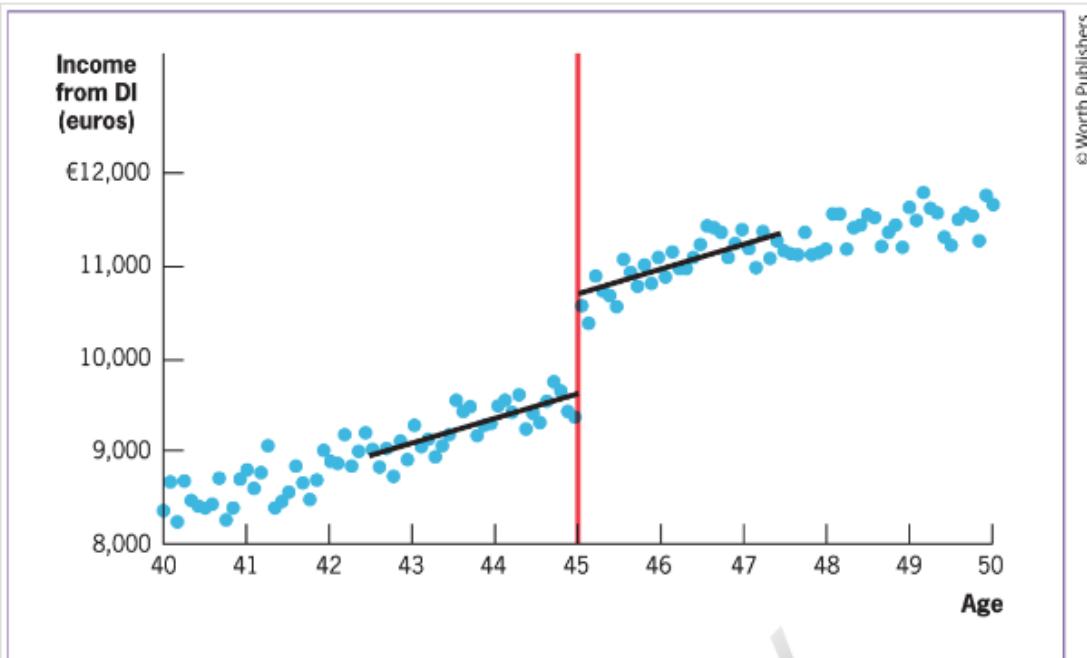


FIGURE 14-5 DI Benefit Amount by Age • As people grow older, they are more likely to receive disability insurance benefits, so that their expected benefit amount rises. But at age 45, there is a jump in benefit amounts that is equivalent to about 1,100 euros. This is likely due to looser screening standards for those older than 45.

Data from: [Borghans et al. \(2014\)](#).

© Worth Publishers

i

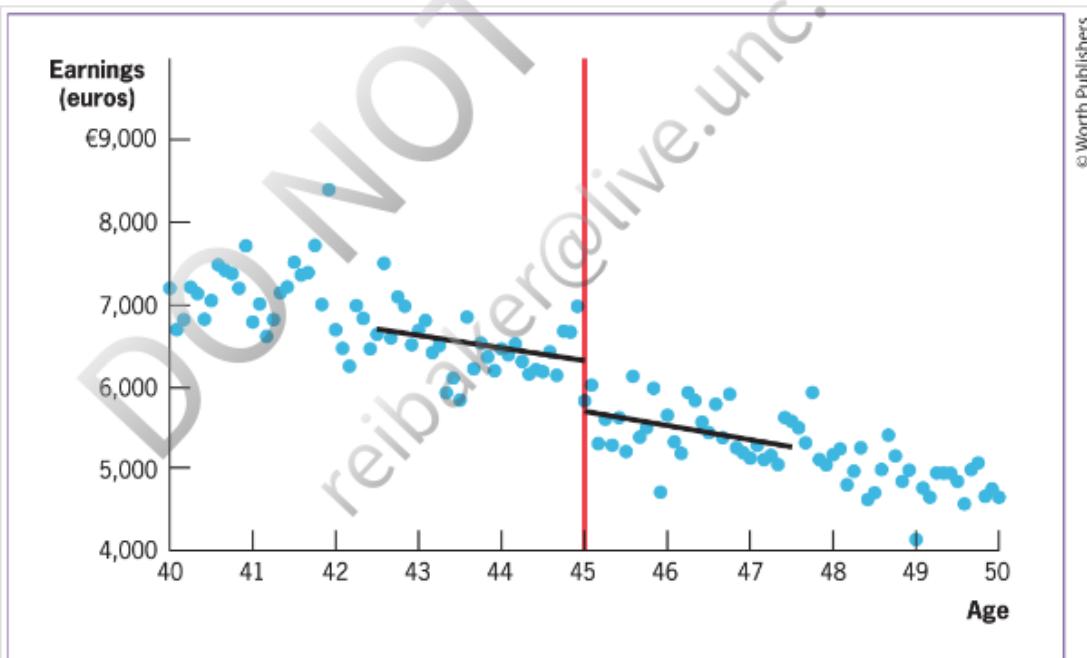


FIGURE 14-6 Earnings by Age • Earnings before the age cutoff are €624 per year higher at the cutoff age than after. This suggests that labor income is crowded out by DI income, which is more easily accessible after turning 45.

Data from: [Borghans et al. \(2014\)](#).

© Worth Publishers

a



[Figure 14-5](#) shows that DI benefits rose with age, as would be expected from rising disability, but that there was a discrete jump at age 45: those over age 45 had benefits that were about 1,100 euros higher, consistent with the looser screening that was in place at that age. [Figure 14-6](#) shows that there was an exactly corresponding decline in earnings at age 45; there is a discrete reduction in the slope of the line to the right of age 45 compared to the left. The results imply a substantial rebound in earnings for those who lost their DI benefits: their earnings rose by 62 euros for each 100 euros in reduced DI benefits.

Taken together, these studies suggest that there is significant moral hazard in DI: many of those relying on DI could be working if the program were not available.

A final test of moral hazard, which most closely parallels work in UI and DI, is to assess how changes in the generosity of DI benefits affect work effort. The most recent available evidence on this topic found that labor supply does respond to benefits generosity, but fairly modestly, with an elasticity of labor supply with respect to benefits of 0.3 to 0.5.³⁰ This is much lower, for example, than Meyer's estimate of an elasticity of 0.8 for unemployment durations with respect to UI benefits. This suggests that at least along the dimension of benefits responsiveness, moral hazard is less of an issue in the DI program than it is in the UI program.

Evidence for Moral Hazard in WC

There is much evidence that points to a major moral hazard effect of the WC program. The first is [Krueger's \(1990\)](#) study of the incidence of injuries, which asked: Does the rate of reported injury on the job depend on the generosity of the WC system? Krueger answered this question by using the fact that benefits for WC vary across the states and looking at how reported rates of on-the-job injury change when states increase their WC benefits. In the absence of moral hazard, there would be no reason for the injury rate to rise just because a state increased its benefits. In fact, however, that is exactly what Krueger found: every 10% increase in benefits generosity led to a 7% rise in the rate of reported injury (a large elasticity of 0.7).

The second piece of evidence comes from [Krueger's \(1991\)](#) study of the impact of WC benefits on injury durations, which is similar to earlier studies of the impact of UI on unemployment durations. In this study, reviewed in the Empirical Evidence box, Krueger found enormous impacts of benefits on injury durations, with each 10% rise in benefits leading to durations that were 17% longer! This is a much larger response than for UI, where 10% higher benefits led to 8% longer unemployment durations.

Additional evidence for moral hazard in WC comes from the *types of injuries* reported. Recall that the problem of moral hazard is a problem of *unobservability* of true injury status. Unobservability implies that moral hazard would be worse for injuries that are hard to observe or verify, such as sprains or strains, and less of a problem for verifiable injuries, such as lacerations or broken or missing limbs. In fact, [Krueger \(1991\)](#) found that the response of injury durations to benefits increases was much stronger for hard-to-verify injuries than for easier-to-verify injuries. That is, a rise in benefits caused a very large increase in reported back sprains, for example, but a smaller increase (or none at all) for lacerations.

A final interesting piece of evidence for the moral hazard effects of WC comes from the infamous “Monday effect.” By examining the types of claims made for WC by day of the week, one sees that on Mondays, there is a large rise in sprains and strains compared to lacerations.³¹ This suggests that many of the reported injuries on Mondays may actually arise from injuries incurred over the weekend and then claimed on Monday in order to qualify for WC. If you strain yourself riding your bike on Sunday, it may be easy to pretend that it happened at work on Monday; but if you cut yourself with a power saw on Sunday, it will be much harder to pretend that this was a work accident on Monday!

EMPIRICAL EVIDENCE

Moral Hazard Effects of Workers' Compensation

[Figure 14-7](#) illustrates the excellent quasi-experiment for the state of Minnesota used by [Krueger \(1991\)](#). Minnesota had a WC program with three flat rates, along with two sloped segments connecting these flat rates. On October 1, 1986, the state increased the benefits along each of the flat-rate portions but left the slopes unchanged. This policy change set up three treatment groups, the workers on each of the flat-rate portions (groups A, C, and E), and two control groups, the workers on the sloped portions (groups B and D). It seems unlikely that there were other factors changing that affected only the three treatment groups and not the two control groups, except for the benefits change that affected the treatments but not the controls. Thus, the treatment and control groups should be comparable; this condition made it possible for Krueger to estimate a causal effect of the benefits change.

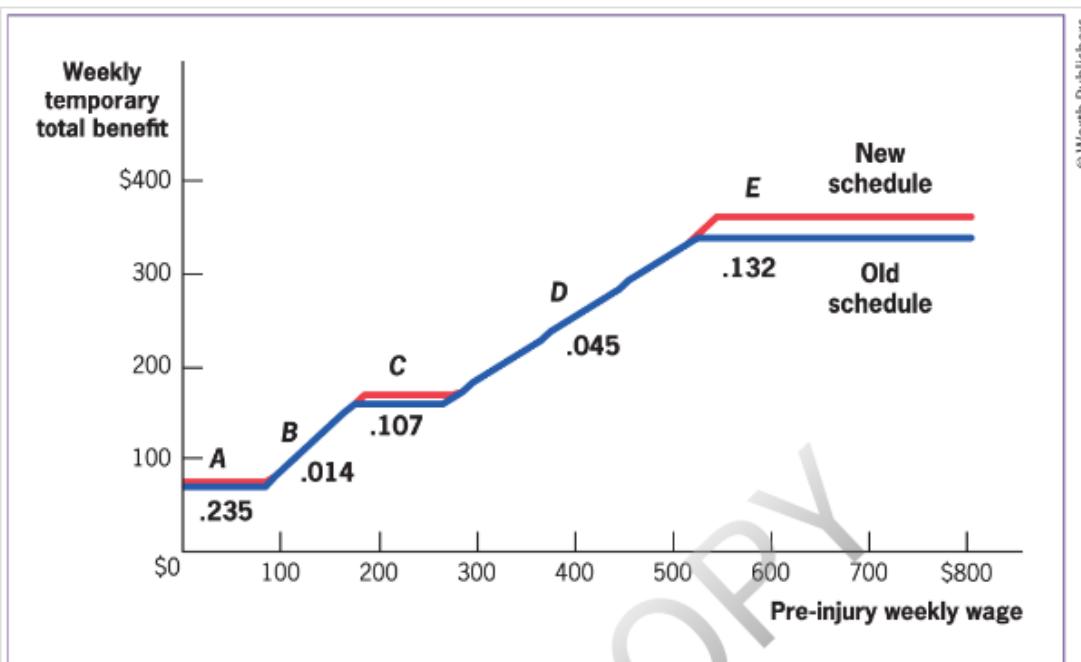


FIGURE 14-7 WC Benefits Changes and Injury Duration • On October 1, 1986, the state of Minnesota raised benefits for workers on segments A, C, and E of this schedule, while leaving benefits unchanged for segments B and D. Krueger found a sizeable rise in the duration of workplace injuries for workers at segments A, C, and E, but not for workers at segments B and D. Numbers below segments of the curve are differences [in log (weeks of benefits received)] between workers injured before and after the benefits change.

Data from: [Krueger \(1991\)](#), Figure 2.

i

The numbers along each segment of this schedule show the percentage change in injury durations from before October 1, 1986, to after that date. For groups A, C, and E (for whom benefits increased), there were large rises in injury durations: 23.5% for group A, 10.7% for group C, and 13.2% for group E. For groups B and D (for whom benefits did not increase), however, the rise in injury durations was negligible (1.4% for group B and 4.5% for group D).³² Thus, increases in benefits appear to be associated with increases in injury duration. The estimated response of injury duration to increased benefits is enormous, with an implied elasticity of 1.7; that is, each 10% rise in benefits led to an injury duration that was 17% longer. Recent research by [Cabral and Dillender \(2020\)](#) finds a sizable but smaller elasticity of 0.7, comparable to the estimates for UI.

A final interesting piece of evidence for the moral hazard effects of WC comes from the infamous “Monday effect.” By examining the types of claims made for WC by day of the week, one sees that on Mondays, there is a large rise in sprains and strains compared to lacerations.³¹ This suggests that many of the reported injuries on Mondays may actually arise from injuries incurred over the weekend and then claimed on Monday in order to qualify for WC. If you strain yourself riding your bike on Sunday, it may be easy to pretend that it happened at work on Monday; but if you cut yourself with a power saw on Sunday, it will be much harder to pretend that this was a work accident on Monday!

EMPIRICAL EVIDENCE

Moral Hazard Effects of Workers' Compensation

Figure 14-7 illustrates the excellent quasi-experiment for the state of Minnesota used by Krueger (1991). Minnesota had a WC program with three flat rates, along with two sloped segments connecting these flat rates. On October 1, 1986, the state increased the benefits along each of the flat-rate portions but left the slopes unchanged. This policy change set up three treatment groups, the workers on each of the flat-rate portions (groups A, C, and E), and two control groups, the workers on the sloped portions (groups B and D). It seems unlikely that there were other factors changing that affected only the three treatment groups and not the two control groups, except for the benefits change that affected the treatments but not the controls. Thus, the treatment and control groups should be comparable; this condition made it possible for Krueger to estimate a causal effect of the benefits change.

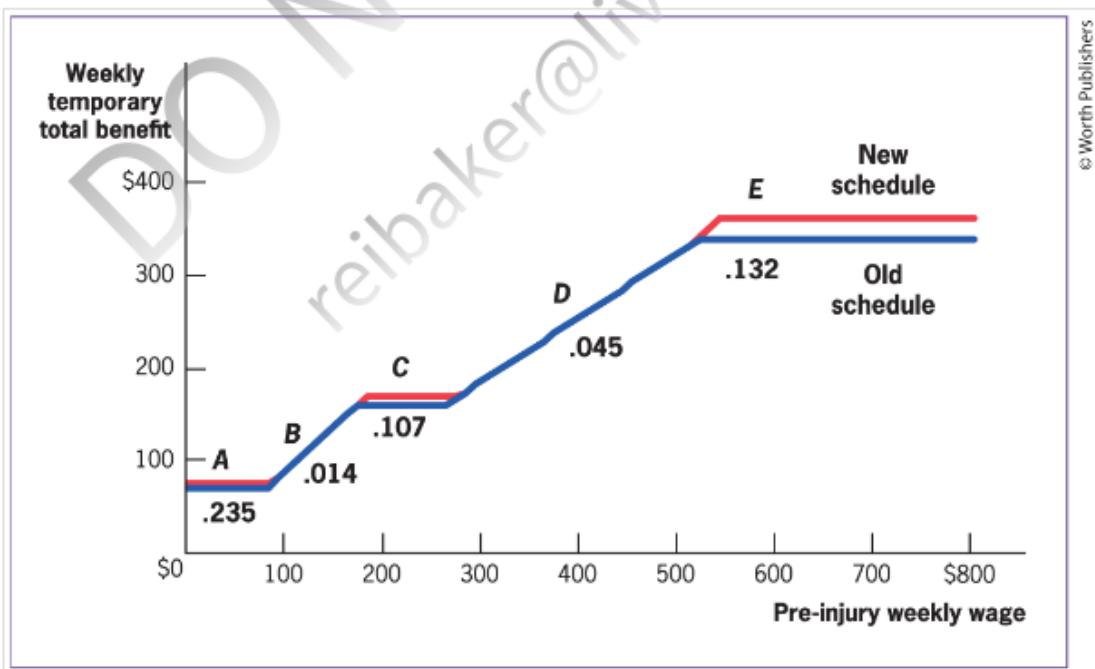


FIGURE 14-7 WC Benefits Changes and Injury Duration • On October 1, 1986, the state of Minnesota raised benefits for workers on segments A, C, and E of this schedule, while leaving benefits unchanged for

segments *B* and *D*. Krueger found a sizeable rise in the duration of workplace injuries for workers at segments *A*, *C*, and *E*, but not for workers at segments *B* and *D*. Numbers below segments of the curve are differences [in log (weeks of benefits received)] between workers injured before and after the benefits change.

Data from: [Krueger \(1991\)](#), Figure 2.

i

The numbers along each segment of this schedule show the percentage change in injury durations from before October 1, 1986, to after that date. For groups *A*, *C*, and *E* (for whom benefits increased), there were large rises in injury durations: 23.5% for group *A*, 10.7% for group *C*, and 13.2% for group *E*. For groups *B* and *D* (for whom benefits did not increase), however, the rise in injury durations was negligible (1.4% for group *B* and 4.5% for group *D*).³² Thus, increases in benefits appear to be associated with increases in injury duration. The estimated response of injury duration to increased benefits is enormous, with an implied elasticity of 1.7; that is, each 10% rise in benefits led to an injury duration that was 17% longer. Recent research by [Cabral and Dillender \(2020\)](#) finds a sizable but smaller elasticity of 0.7, comparable to the estimates for UI.

14.4 The Costs and Benefits of Social Insurance to Firms

Thus far, our analysis has focused on the impact of the social insurance programs on workers. Yet because employers play an important role in both layoffs and work injuries, both UI and WC programs can have important effects on firms' decision making as well. We first discuss the impact of UI on firms and then turn to the impact of WC.

The Effects of Partial Experience Rating in UI on Layoffs

The key feature of UI from the firm's perspective is that it is *partially experience-rated*. A fully experience-rated system would be one in which a firm pays additional tax each time it lays off a worker. The amount collected through the tax would equal the expected UI benefits paid to that worker. In the current UI system, payroll taxes do rise with past layoffs, but much less than one for one.

The degree of partial experience rating is illustrated in [Figure 14-8](#), which shows the relationship between the UI payroll tax rate and past layoffs in Vermont, which has a typical state experience-rating system. There is a minimum payroll tax rate of 1.1% paid by all firms in the state. Beyond that point, the payroll tax rate rises with what is called the *benefit ratio*, the ratio of the payments made by the UI system to the firm's laid-off workers, relative to the size of the firm's payroll, averaged over the past four years. Thus, a benefit ratio of 10 means that UI paid benefits equal to 10% of the firm's payroll on average over the past four years. Once this measure of past layoffs reaches a particular level, the payroll tax rate climbs, continuing to ratchet upward in this fashion until it reaches a maximum rate of 7.7% for firms with a benefits ratio of 20 or higher.

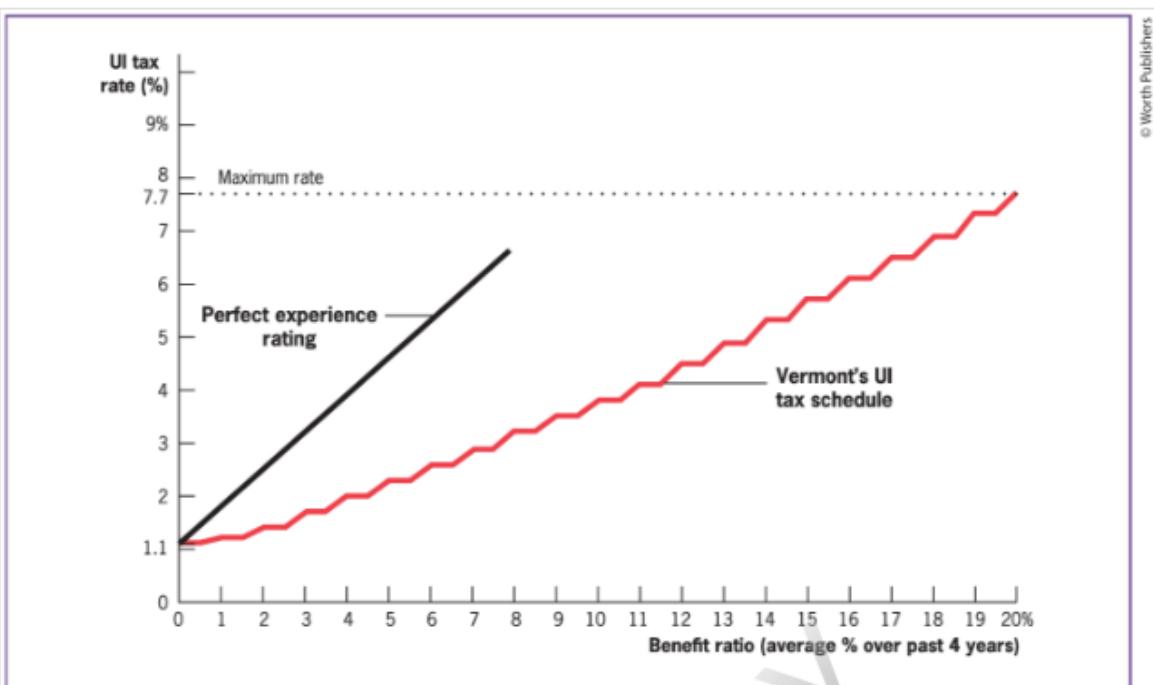


FIGURE 14-8 Experience-Rating Schedule for Vermont • In Vermont, as in most states, the UI tax rate paid by employers rises as past layoffs rise, as measured by the benefit ratio, which is the ratio of UI benefits paid to the firm's workers relative to the firm's payroll. This rise is not on the one-for-one basis, however, that would follow the perfect experience-rating line. As a result, high-layoff (high-benefit ratio) employers are relatively subsidized by the system.

Data from: Tax rates taken from the Vermont Department of [Labor \(2018\)](#).



Partial Experience-Rating Subsidizes Layoffs

A fully experience-rated system would follow the path of a 45-degree line from the origin, which would indicate a one-for-one increase in tax payments for each increase in benefits payouts. Relative to full experience rating, the Vermont system causes firms with a low level of layoffs to pay too much (because the actual schedule is above the 45-degree line), and firms with many layoffs to pay too little (because the actual schedule is below the 45-degree line). Thus, relative to a full system of experience rating, partial experience rating subsidizes firms with high layoff rates. This conclusion is best illustrated by the maximum condition: once firms reach the maximum UI tax rate, there is no additional tax cost to them for additional layoffs. This is a general phenomenon in states across the nation: high-layoff firms are subsidized for additional layoffs, relative to a fully experience-rated system.

How is this a subsidy? Think about the firm and the worker as making a joint decision about whether to place the worker on *temporary layoff*, where the worker spends some time laid off but with a promise (explicit or implicit) to be hired back

after the layoff is over (like my wife's aunt in the example in [Chapter 12](#)). From the worker's perspective, a temporary layoff is some time off at a partial wage—a partially paid vacation. From the firm's perspective, the attractiveness of a temporary layoff depends on the extent of experience rating. If there is no experience rating, the firm pays nothing when a worker is temporarily laid off. So this is a vacation for workers, paid for by the government—a good deal for both firms and workers.

Contrast this with full experience rating. In that case, the benefits paid to the worker would be exactly canceled by the taxes paid by the firm so that, on net, there would be no money flowing from the government to this worker-firm pair. Thus, if the worker wants a paid vacation, the firm has to pay for it; there is no subsidy to layoffs from the government because the firm pays the full costs of any layoffs.

Evidence on Effect of Partial Experience Rating on Layoffs

It is clear from the previous discussion that firms should have more layoffs when a state's experience-rating system is more partial (i.e., the more that the government, and not the firm, pays for benefits when the worker is laid off). In fact, there are several careful investigations of the impact of the experience-rating structure on layoff decisions. These studies investigated state systems with different degrees of experience rating and found that partial experience rating increases the rate of temporary layoffs, as predicted by the theory. The studies suggest that partial experience rating alone can account for 21 to 33% of all temporary layoffs in the United States.³³

The “Benefits” of Partial Experience Rating

If partial experience-rating systems increase the number of layoffs in the U.S. labor market, why are they so common in state UI programs? Once again, the benefit that offsets this moral hazard cost is consumption smoothing. Having a fully experience-rated system would “hit firms while they are down”: just when firms have laid off the most workers (presumably because the firm is not doing well), their taxes would increase the most.

This may be efficient if this provides protection to risky, high-growth firms; [Van Doornik et al. \(2020\)](#) find that a non-experience rated system in Brazil redistributed from safe to risky firms. But it may be inefficient if these are firms that should just go out of business leaving the field to their more efficient rivals. In particular the

go out of business, leaving the field to their more efficient rivals. In particular, the system of partial experience rating subsidizes seasonal firms, which can afford to hire workers for just part of the year because the workers can receive UI for the remainder of the year, at no marginal cost to the firm. [Anderson and Meyer \(2000\)](#) computed that, in Washington State, the one-eighth of firms that are subsidized by the UI system for four continuous years account for one-third of all UI benefits payments.

APPLICATION

The “Cash Cow” of Partial Experience Rating



The United States is actually relatively unique among industrialized nations in having *any* experience rating of its UI system, however partial. Most nations finance their UI systems through a flat payroll tax that is unrelated to the firm’s actual layoff experience. Such a system can turn the UI system into a “cash cow” that subsidizes the existence of highly inefficient firms.

Consider the case of Canada’s UI system, where workers traditionally had to work only 10 weeks to qualify for 42 weeks of UI with a replacement rate of 60%, at no extra cost to their firms. One can assess the implications of such a generous system by posing a hypothetical example. You and four friends want to figure out a way to work only 10 weeks a year and take vacation the rest of the year. The five of you are considering buying a fishing boat, with each of you working 10 weeks out of the year, for a total of 50 weeks. You know that with this fishing boat, you can catch \$40,000 worth of fish during the entire year. That is only \$8,000 of fish for each 10-week period, a sum that is not large enough to support each one of you for an entire year. In the absence of UI, then you would not purchase this fishing boat, which is the socially efficient outcome: it would be inefficient for your combined group of workers to produce only \$40,000 worth of goods per year.

Given the structure of the Canadian UI system, however, this purchase will be more attractive. In this system, you and your friends each work 10 weeks, for a total of 50 workweeks. Each of you would report earning \$800 per week for the 10 weeks worked and then report being laid off. Because you report being laid off, you would each receive \$20,160 of UI during the rest of the year (60% of the \$800 per week reported earnings, for 42 weeks). So the total UI benefits income across all five recipients is $\$20,160 \times 5 = \$100,800$. Adding the \$40,000 of fish produced, that is a total income to the five of you of \$140,800, or \$28,160 per worker per year for only 10 weeks of work each! This makes the purchase of the fishing boat much more attractive.

This example illustrates the fundamental problem with partial experience rating: it subsidizes the existence of inefficient firms. Your firm is not an economically viable employer of five employees. It could pay each employee \$8,000 per year if there were no UI system. But by exploiting partial experience rating, the firm can remain viable, with each employee earning \$28,160 for the year. Thus, UI is not simply a system of insurance against true unemployment risk in Canada, but also a large government transfer to inefficient firms and their laid-off workers. ■

Workers' Compensation and Firms

A similar set of issues arises in the context of workers' compensation. With UI, partial experience rating means that firms and workers have an incentive to increase layoffs to exploit this government payment for leisure. With WC, firms and workers can get together to increase "injuries" if the insurance is less than fully experience-rated. There is an additional problem as well: firms have less incentive to invest in safety when there is no-fault insurance for injuries. In the past, when injuries could lead to lawsuits against a firm, firms had to trade off the cost of making workplaces safer against the costs of getting sued if someone was injured. Now, for the partially experience-rated firm, there are relatively little savings to making the workplace safer because the firm pays relatively little of the cost of the WC benefit if the worker is injured. Thus, an additional moral hazard for firms is that no-fault WC can lead to less-safe workplaces.

The potentially important role played by experience rating in WC is illustrated in [Krueger's \(1991\)](#) study. The study compares the injury durations of employees of firms that *self-insure* their WC costs to the durations of employees of firms that buy their WC in a partially experience-rated insurance market. Self-insured firms pay their own injury claims, so by definition, they are fully experience-rated (more injuries lead to higher costs for the firm). Krueger found that workers in self-insured firms came back to work more quickly from a given type of injury than did workers in comparable non-self-insured (partially experience-rated) firms. Moreover, in self-insured firms, worker injury durations were much less responsive to benefits increases than durations in non-self-insured firms. These findings suggest that firms that are more fully experience-rated are much more aggressive in monitoring worker injury durations.

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DONOTCOPY
reibaker@live.unc.edu

Consider the case of Canada's UI system, where workers traditionally had to work only 10 weeks to qualify for 42 weeks of UI with a replacement rate of 60%, at no extra cost to their firms. One can assess the implications of such a generous system by posing a hypothetical example. You and four friends want to figure out a way to work only 10 weeks a year and take vacation the rest of the year. The five of you are considering buying a fishing boat, with each of you working 10 weeks out of the year, for a total of 50 weeks. You know that with this fishing boat, you can catch \$40,000 worth of fish during the entire year. That is only \$8,000 of fish for each 10-week period, a sum that is not large enough to support each one of you for an entire year. In the absence of UI, then you would not purchase this fishing boat, which is the socially efficient outcome: it would be inefficient for your combined group of workers to produce only \$40,000 worth of goods per year.

Given the structure of the Canadian UI system, however, this purchase will be more attractive. In this system, you and your friends each work 10 weeks, for a total of 50 workweeks. Each of you would report earning \$800 per week for the 10 weeks worked and then report being laid off. Because you report being laid off, you would each receive \$20,160 of UI during the rest of the year (60% of the \$800 per week reported earnings, for 42 weeks). So the total UI benefits income across all five recipients is $\$20,160 \times 5 = \$100,800$. Adding the \$40,000 of fish produced, that is a total income to the five of you of \$140,800, or \$28,160 per worker per year for only 10 weeks of work each! This makes the purchase of the fishing boat much more attractive.

This example illustrates the fundamental problem with partial experience rating: it subsidizes the existence of inefficient firms. Your firm is not an economically viable employer of five employees. It could pay each employee \$8,000 per year if there were no UI system. But by exploiting partial experience rating, the firm can remain viable, with each employee earning \$28,160 for the year. Thus, UI is not simply a system of insurance against true unemployment risk in Canada, but also a large government transfer to inefficient firms and their laid-off workers. ■

Workers' Compensation and Firms

A similar set of issues arises in the context of workers' compensation. With UI, partial experience rating means that firms and workers have an incentive to increase layoffs to exploit this government payment for leisure. With WC, firms and workers can get together to increase "injuries" if the insurance is less than fully experience-rated. There is an additional problem as well: firms have less incentive

to invest in safety when there is no-fault insurance for injuries. In the past, when injuries could lead to lawsuits against a firm, firms had to trade off the cost of making workplaces safer against the costs of getting sued if someone was injured. Now, for the partially experience-rated firm, there are relatively little savings to making the workplace safer because the firm pays relatively little of the cost of the WC benefit if the worker is injured. Thus, an additional moral hazard for firms is that no-fault WC can lead to less-safe workplaces.

The potentially important role played by experience rating in WC is illustrated in [Krueger's \(1991\)](#) study. The study compares the injury durations of employees of firms that *self-insure* their WC costs to the durations of employees of firms that buy their WC in a partially experience-rated insurance market. Self-insured firms pay their own injury claims, so by definition, they are fully experience-rated (more injuries lead to higher costs for the firm). Krueger found that workers in self-insured firms came back to work more quickly from a given type of injury than did workers in comparable non-self-insured (partially experience-rated) firms. Moreover, in self-insured firms, worker injury durations were much less responsive to benefits increases than durations in non-self-insured firms. These findings suggest that firms that are more fully experience-rated are much more aggressive in monitoring worker injury durations.

14.5 Implications for Program Reform

This chapter has presented a large body of evidence on the costs and benefits of three of the most important social insurance programs in the United States. We can use this evidence, along with the theory presented in [Chapter 12](#), to draw lessons for program reform.

Benefits Generosity

The optimal level of benefits generosity reflects the trade-off between moral hazard and consumption-smoothing benefits. It is clear that for all three programs studied here, the replacement rate should be less than 100% because there is significant moral hazard associated with each type of insurance. The literature also indicates that the negative behavioral responses to these programs (such as longer unemployment or injury durations) are very large for WC, fairly large for UI, and smaller for DI. At the same time, the consumption-smoothing benefits are likely largest for DI and have been shown to be only partial for UI (and likely for WC as well). Taken together, these facts suggest that benefits should be highest for DI and lowest for WC, with UI in the middle.

As [Table 14-2](#) shows, however, this is not the case. In fact, WC has the most generous benefits of all of these programs, which is clearly inconsistent with the evidence presented here.

Targeting

Another issue that is raised by the discussions of [Chapter 12](#) and the evidence in this chapter is the need to better target program benefits toward those who benefit the most from consumption smoothing and/or for whom the moral hazard problems of social insurance are smallest. Consider those who regularly have temporary unemployment spells and receive implicit promises from their employers that they can return to their old jobs. For this group, there is little consumption-smoothing benefit of UI because the predictability and regularity of such layoffs should allow them to use self-insurance to smooth consumption. Moreover, empirical analyses have shown that this is the group that is the most responsive to UI benefit levels in terms of extending unemployment durations. For this group, the costs of high UI benefits appear to outweigh their benefits. Thus, efficiency could be improved if UI benefits could be targeted away from this group and toward those who have been permanently laid off.

Targeting is also possible within the DI and WC programs, based on the type of injury or disability. Some injuries or disabilities are easier to diagnose, minimizing problems of moral hazard. Becoming blind or paralyzed, or having a laceration or losing a limb, is unlikely to represent a negative behavioral response to social insurance program generosity. In principle, it would be possible to arrange these programs so that higher benefits were paid to people with less ambiguous (easier to verify) disabilities or injuries, people for whom the consumption-smoothing benefits are more likely to outweigh the moral hazard costs. Indeed, as noted earlier, such a system is in place in the Netherlands and other nations. And both [French and Song \(2014\)](#) and [Borghans et al. \(2014\)](#) find that the moral hazard effects of DI are smaller for younger (French and Song) or less disabled (Borghans et al.) workers. Moreover, [Meyer and Mok \(2018\)](#) find much larger consumption drops for those with more severe disabilities. In practice, however, this approach raises difficult issues of how to classify injuries into these different categories.³⁴

Worker Self-Insurance?

A more radical reform of the three social insurance systems would be to move toward *worker self-insurance* against these adverse events. For example, the government could replace payroll taxes and mandated WC insurance with individual “social insurance savings accounts,” to which workers would contribute some fixed amount. If they qualified for social insurance because they experienced one of these adverse events, they could then draw on this savings account, with provisions for borrowing (and repaying) if they exceed the account. If there are positive balances in these accounts at retirement, they could be used to finance retirement consumption.

A system of worker self-insurance for UI would be similar to a privatized Social Security system, and it has many of the same advantages and disadvantages. One major advantage in this context is that there is much clearer evidence for moral hazard effects of programs such as UI and WC than for Social Security and likely much less consumption smoothing provided by government insurance. By making unemployed or injured workers pay for their income support out of their own savings accounts, the program would minimize moral hazard. [Feldstein and Altman \(2007\)](#) concluded that a UI payroll tax of 4% invested in such accounts could cover the costs of unemployment spells for virtually all workers. But a disadvantage of such a self-insured system is that we would lose the redistribution of income from those who have not lost their jobs to those who have.

APPLICATION

Modernizing UI



Unemployment Insurance was introduced in 1935, but the system has largely remained unchanged for decades. The explosion of UI claims during the Covid-19 crisis exposed a multitude of underlying weaknesses in our UI system and has led to wide-ranging discussion about how to modernize our UI system.

DO NOT COPY
reibaker@live.unc.edu

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DONOTCOPY
reibaker@live.unc.edu

Finally, we should consider extending the UI system beyond its narrow focus on income replacement for the unemployed. Most European nations pair their unemployment insurance systems with aggressive job search assistance and placement services, and these are estimated to provide considerable improvements in job placement.⁴⁴ The state of Nevada experimented with personalized job counseling services for a random sample of new UI recipients in 2009 and found that the program led to substantial short-term reductions in UI receipt, as well as to long term-increases in employment and wages.⁴⁵

Another potential reform would be to address the fact that while UI replaces the short-term income loss due to unemployment, it does not help the many displaced workers who see sizeable long-term reductions in their earnings. On average, full-time workers displaced from a job due to economic reasons (such as a plant closing) who then find another full-time job see their earnings decline by 17% relative to comparable workers who do not lose their job ([Farber, 2005](#)). This is a much bigger reduction in lifetime resources than a 26-week temporary loss of earnings, yet there is no social insurance to protect against this loss in earnings. [Kling \(2006\)](#) suggests supplementing UI with a system of reemployment earnings insurance, which would provide a transfer from the government equal to 25% of the difference between the old hourly wage and the new hourly wage. The duration of payments (up to a maximum of six years) would be determined by the number of hours worked in the two years prior to the job loss. Partial replacement, only 25% of the wage difference, ensures that workers will still try to find a high-wage new job. Indeed, a social experiment of a similar type of program in Canada found that it provided significant insurance for wage losses without adverse effects on job search or work behavior. ■

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DONOTCOPY
reibaker@live.unc.edu

Finally, we should consider extending the UI system beyond its narrow focus on income replacement for the unemployed. Most European nations pair their unemployment insurance systems with aggressive job search assistance and placement services, and these are estimated to provide considerable improvements in job placement.⁴⁴ The state of Nevada experimented with personalized job counseling services for a random sample of new UI recipients in 2009 and found that the program led to substantial short-term reductions in UI receipt, as well as to long term-increases in employment and wages.⁴⁵

Another potential reform would be to address the fact that while UI replaces the short-term income loss due to unemployment, it does not help the many displaced workers who see sizeable long-term reductions in their earnings. On average, full-time workers displaced from a job due to economic reasons (such as a plant closing) who then find another full-time job see their earnings decline by 17% relative to comparable workers who do not lose their job ([Farber, 2005](#)). This is a much bigger reduction in lifetime resources than a 26-week temporary loss of earnings, yet there is no social insurance to protect against this loss in earnings. [Kling \(2006\)](#) suggests supplementing UI with a system of reemployment earnings insurance, which would provide a transfer from the government equal to 25% of the difference between the old hourly wage and the new hourly wage. The duration of payments (up to a maximum of six years) would be determined by the number of hours worked in the two years prior to the job loss. Partial replacement, only 25% of the wage difference, ensures that workers will still try to find a high-wage new job. Indeed, a social experiment of a similar type of program in Canada found that it provided significant insurance for wage losses without adverse effects on job search or work behavior. ■

14.6 Conclusion

The three social insurance programs studied in this chapter (unemployment insurance, disability insurance, and workers' compensation) provide excellent applications of the general principles of social insurance covered in [Chapter 12](#). In each case, individuals clearly value the consumption smoothing provided by the program, and in each case, there are significant moral hazard costs associated with the provision of the insurance. These moral hazard costs dictate that insurance be less than full. The many empirical analyses of all three programs can be used to inform policy makers' decisions as program reforms move forward.

HIGHLIGHTS

- Unemployment insurance (UI) typically provides 26 weeks of benefits to workers who are laid off and searching for work.
- Disability insurance (DI) provides income replacement at Social Security levels for workers who are disabled and can no longer work.
- Workers' compensation (WC) provides generous cash benefits and medical insurance to workers injured on the job.
- The consumption smoothing provided by UI is only partial because the program seems to a large extent to crowd out other forms of self-insurance. The effects of DI and WC on consumption smoothing have not been studied enough to reach any conclusion about their consumption-smoothing effects.
- Existing evidence suggests that the moral hazard costs of UI are large, with an elasticity of unemployment durations with respect to benefits of 0.8. (Each 10% rise in benefits raises unemployment durations by 8%.)
- There are also significant moral hazard effects associated with DI, although they are smaller, with each 10% rise in benefits leading to only a 3-5% rise in the odds of leaving the labor force.
- The largest moral hazard effects appear to be associated with WC, with each 10% rise in benefits leading to a 7% rise in the odds of claiming a workplace injury, and leading to a 17% rise in injury durations.
- Partial experience rating of both the UI and WC programs appears to raise the rate of layoffs and the duration of injury-related job leave.
- The evidence in this chapter implies that the WC program is likely too generous, and the DI program perhaps not generous enough; that benefits should be targeted toward groups such as the long-term unemployed and seriously disabled; and that firms should be more fully experience-rated for social insurance payments to their workers.

QUESTIONS AND PROBLEMS

1. The UI payroll tax is said to be partially experience-rated because the tax rate on earnings is higher for firms with a history of laying off workers. What is the rationale for making the payroll tax rate a function of a firm's layoff history?
2. Describe the effects of raising the maximum benefit level for UI on the savings rate of high-income workers. How big are the consumption-

- smoothing benefits of this policy change likely to be? Are there other potential benefits of raising this maximum benefit level?
3. **E** Workers' compensation benefits vary across states and types of injuries. How can you employ this information to estimate the elasticity of injury with respect to workers' compensation benefits generosity?
 4. The Organization for Economic Cooperation and Development (OECD) creates tax-benefit models to estimate how tax and benefit rules in OECD countries affect the net income of workers when they move between unemployment and employment. OECD tax-benefit data portal, currently found at <https://www.oecd.org/els/soc/benefits-and-wages/data/>, and click on "Net replacement rates in unemployment," a web data set portal currently found at <https://stats.oecd.org/Index.aspx?DataSetCode=NRR>.
 - a. Look at the two data series "Net Replacement Rate in unemployment" and "Historical Gross Replacement Rates of unemployment benefits." Read the explanations in the Information tabs to understand the differences between two similar statistics: the net replacement rate (NRR) and the gross replacement rate (GRR). Which of the two do you think does a better job of capturing the most relevant factors that unemployed individuals usually consider as they decide whether to reenter the work force?
 - b. For Net Replacement Rate in unemployment (NRR), compare the NRR for years 2001 and 2019. Consider a single person without children unemployed for two months with an average wage for their previous in-work earnings and include in the calculation social assistance and housing benefits. Contrast this single individual to a couple with two children where one partner is out of work. Between 2001 and 2019, which three countries experienced the largest drops in NRR for the individual? The couple? What about the three countries with the largest increases? Do benefit levels seem to be moving similarly for individual single workers and families?
 5. What does the empirical evidence on the consumption-smoothing benefits of UI indicate about the degree to which individuals are, on average, insured against the income losses associated with unemployment?
 6. **E** Consider Meyer's 1989 study of the effects of unemployment benefits on unemployment spell durations. How does this study deal with the likelihood that unemployment spells and unemployment benefits may both increase during economic recessions?
 7. Gruber (2000) found evidence that the elasticity of labor supply with respect to DI benefits is considerably smaller than the estimates of the

The expected duration is considerably smaller than the estimates of the

elasticity of unemployment durations with respect to UI benefits. Why might moral hazard be less of an issue in the DI program than in the UI program?

8. Governments typically provide DI and UI to workers. In contrast, governments typically *mandate* that firms provide workers' compensation insurance to their workers but do not provide the coverage. Why the difference? Why don't governments provide workers' compensation instead of mandating it?
9. In May 2004, the state of Vermont significantly reformed its workers' compensation system. One key provision of this reform was to reduce the window of time during which a claimant could file an initial workers' compensation claim. Might this reform have helped to reduce the degree of fraudulent use of the workers' compensation system? Explain.
10. Senator Doppelganger has proposed rules that will make it easier for workers to apply for and receive disability benefits. What is this likely to do to rates of application for disability benefits? To the reported unemployment rate?
11. Unemployment benefits are subject to income taxes. They are not (by default) subject to withholding or to payroll taxes. Suppose that a particular worker is subject to a payroll tax of 7.65% of their wages; that they pay an average income tax rate of 10%; and that after losing their job, they are eligible to receive 60% of her previous weekly income.
 - a. What fraction of their take-home paycheck will they receive from unemployment insurance?
 - b. What effect does the tax treatment of unemployment insurance benefits have on the ability it provides its recipients for smoothing?

ADVANCED QUESTIONS

12. Are individuals more likely to maintain their pre-injury consumption levels after an easily preventable on-the-job injury than after a difficult-to-prevent on-the-job injury? Explain.
13. **E** The empirical evidence on unemployment spell durations suggests that workers who leave unemployment earlier (i.e., find or take a job sooner) have no higher post-unemployment wages than workers who leave unemployment later. This result could be interpreted as evidence

that the quality of the job match does not improve as the unemployment spell grows longer.

- a. What does this interpretation of the evidence imply about the moral hazard costs of UI?
 - b. An alternative explanation for this evidence is that workers with longer unemployment spells are less qualified than are workers with shorter unemployment spells. How could you empirically distinguish between this explanation and the explanation put forth in (a)?
14. Did the changes made to the U.S. unemployment insurance system during the Covid-19 pandemic increase or decrease the consumption-smoothing benefits of the system compared to the years prior to the pandemic?
15. You are hired by the presidential administration to review the UI program, which currently replaces approximately 45% of a worker's wages for 26 weeks after they lose their job. Consider two alternative reforms of the current UI system. The first is to experience-rate firms fully, so that the taxes firms pay are set exactly equal to the benefits their workers receive (benefits remain at 45% of wages). The second is a system of *individual* full experience rating—the government lends individuals 45% of their wages while unemployed, but the individuals have to pay this back when they get new jobs.
- a. Contrast the effects of these alternative policies on unemployment durations and the likelihood of worker layoffs.
 - b. What are the consumption-smoothing properties of each alternative policy?

The  icon indicates a question that requires students to apply the empirical economics principles discussed in [Chapter 3](#) and the Empirical Evidence boxes.
arbuz/Shutterstock

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
reibaker@live.unc.edu

Are individuals more likely to maintain their pre-injury consumption levels after an easily preventable on-the-job injury than after a difficult-to-prevent on-the-job injury? Explain.

-  The empirical evidence on unemployment spell durations suggests that workers who leave unemployment earlier (i.e., find or take a job sooner) have no higher post-unemployment wages than workers who leave unemployment later. This result could be interpreted as evidence that the quality of the job match does not improve as the unemployment spell grows longer.
 - a. What does this interpretation of the evidence imply about the moral hazard costs of UI?
 - b. An alternative explanation for this evidence is that workers with longer unemployment spells are less qualified than are workers with shorter unemployment spells. How could you empirically distinguish between this explanation and the explanation put forth in (a)?
- Did the changes made to the U.S. unemployment insurance system during the Covid-19 pandemic increase or decrease the consumption-smoothing benefits of the system compared to the years prior to the pandemic?
 - a. You are hired by the presidential administration to review the UI program, which currently replaces approximately 45% of a worker's wages for 26 weeks after they lose their job. Consider two alternative reforms of the current UI system. The first is to experience-rate firms fully, so that the taxes firms pay are set exactly equal to the benefits their workers receive (benefits remain at 45% of wages). The second is a system of *individual* full experience rating—the government lends individuals 45% of their wages while unemployed, but the individuals have to pay this back when they get new jobs.
 - a. Contrast the effects of these alternative policies on unemployment durations and the likelihood of worker layoffs.
 - b. What are the consumption-smoothing properties of each alternative policy?

The  icon indicates a question that requires students to apply the empirical economics principles discussed in [Chapter 3](#) and the Empirical Evidence boxes.

arbuz/Shutterstock

APPENDIX TO CHAPTER 14: Advanced Quasi-Experimental Analysis



As discussed in this chapter, Meyer's study of the effect of UI benefits on unemployment durations considers two possible control groups: high-earning workers in other states and low-earning workers in the same state. The chapter describes how these two different control groups can be used to compute two difference-in-difference estimators. This appendix describes how these estimates can be combined to form an even more convincing quasi-experimental estimate (known as a “difference-in-difference-in-difference” estimate) of the impact of UI benefits on unemployment durations.

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
reibaker@live.unc.edu

CHAPTER 15

Health Insurance I: Health Economics and Private Health Insurance



James Andrews/Stock/Getty Images

15.1 An Overview of Health Care in the United States

15.2 How Generous Should Insurance Be to Patients?

15.3 How Generous Should Insurance Be to Medical Providers?

15.4 Conclusion

Questions to keep in mind

- What are the different sources of insurance for Americans?
 - How generously should insurance cover illness?
 - How generously should insurance reimburse medical providers?
-

On March 30, 1981, only two months after entering office, President Ronald Reagan was shot by John Hinckley, a deranged fan of Jodie Foster who thought the actress would pay more attention to him if he killed the president. Reagan was rushed to George Washington Hospital, where he underwent three hours of surgery to remove a bullet that had entered his lung. Though the injury was potentially fatal, Reagan made a full recovery and went on to serve a full eight years in office.

One hundred years earlier, President James Garfield was not so lucky. He, too, was shot two months after entering office, on July 2, 1881, by Charles Guiteau, who thought God had ordered him to kill the president. Garfield was brought back to the White House, where for 80 days, a dozen doctors attempted to find one bullet that, unbeknownst to them, had lodged itself near Garfield's spine. The doctors probed the president's wound with their unsterilized fingers and metal rods, succeeding only in widening the wound, infecting it, and puncturing his liver. Alexander Graham Bell ran a metal detector over Garfield's body and soon announced that he'd found the bullet. Surgeons went to work but still failed to locate it, not realizing that Bell had mistakenly detected one of the bedsprings underneath the president. The infection resulting from his poor medical care soon caused Garfield to have a heart attack, and when he eventually died, his coroner declared that Garfield would have survived if only his doctors had left him alone.¹

Between President Garfield's death and President Reagan's election, the field of health care clearly made great strides, so that doctors are now helping rather than harming people. This improvement in the quality of health care in the United States has been accompanied by an enormous increase in the share of the U.S. economy devoted to health care. In 1950, only 5% of U.S. GDP was accounted for by the health care sector. At that time, Americans spent less on health care than on cars, fuel, or clothing. By 2019, health care accounted for 17.7% of GDP, surpassing spending on housing and food.² This growth is not expected to stop: health care spending is expected to grow at an average annual rate of 5.4% through 2028, outpacing projected GDP growth by 1.1%.³

Is such high and rapidly growing health care spending a problem? After all, what is more important than our health? And, by some measures, we are buying wonderful things with our health care dollars. Consider the treatment of knee injuries in the 1950s and today. Seventy years ago, if you tore the meniscus (the cartilage under the kneecap), the only option was to have open surgery, during which the surgeon cut open your knee and removed the entire meniscus. You would spend days in the hospital, months recovering, and 15 years down the road, you would have an increased chance of developing arthritis in that knee. If you tear your meniscus today, you can often have only a small piece of it removed by arthroscopic surgery, which allows the surgeon to make tiny incisions in your knee to repair the damage in an average of 30 minutes. You go home that same day, can do light work within a few days, and be up and running (or whatever other sport you enjoy) within three to six weeks.

Similarly, in 1950, 6 out of every 1,000 Americans died from a heart attack. Since then, that number has fallen to 1.6 out of 1,000.⁴ In 1950, 29 out of every 1,000 infants born died in their first year of life; today, that figure is 5.6 out of 1,000.⁵

Despite the huge benefits reaped from the U.S. health care system, all is not completely well: there are enormous disparities in medical outcomes. For example, in 2018, the White *infant mortality* rate in the United States, the share of infants who die in their first year, was 4.6, which was in line with other developed nations like Canada and Greece. The 2018 Black infant mortality rate, however, was 10.8, which was nearly as high as the infant mortality rate in Libya (11.5) and twice as high as the rate in Serbia (5.3).⁶ In the 22 of the 55 neighborhoods in Baltimore, Maryland, the life expectancy is below that of North Korea.⁷

These disparities can be traced historically to the fact that the United States is the only major industrialized nation that did not endeavor to provide universal access to health care for its citizens. As a result, almost 50 million persons—more than one-sixth of the U.S. population—were without health insurance in 2010. As we've seen, the Affordable Care Act (ACA) has attempted to address this issue, and as of 2020, more than 16 million Americans have gained coverage, but large gaps in coverage remain.⁸

Despite the perceived "private" nature of the U.S. health care system, governments account for almost half of all health spending in the United States. Health care spending is now a quarter of the federal government budget and a third of state and local government budgets.⁹

Furthermore, growth in health spending is projected to account for most of the long-run fiscal problems faced by the U.S. government because of the aging of the U.S. population and the rapid rise in medical care costs. Thus, there are clear public finance issues raised by this large and growing health economy. In this chapter and the next, we discuss many issues relative to health care and its importance to the economy and to government policy. In this chapter, we discuss the nature of health care and the set of general health insurance issues relevant to government involvement in the delivery of health insurance. This chapter provides the basis for understanding the health economy and allows us to contemplate reforms in the government role in the delivery of health care. In the next chapter, we examine the two largest public-sector interventions in health insurance markets—the Medicaid and Medicare programs—and the implications of past evidence for future directions in health care reform.

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DONOTCOPY
reibaker@live.unc.edu

Furthermore, growth in health spending is projected to account for most of the long-run fiscal problems faced by the U.S. government because of the aging of the U.S. population and the rapid rise in medical care costs. Thus, there are clear public finance issues raised by this large and growing health economy. In this chapter and the next, we discuss many issues relative to health care and its importance to the economy and to government policy. In this chapter, we discuss the nature of health care and the set of general health insurance issues relevant to government involvement in the delivery of health insurance. This chapter provides the basis for understanding the health economy and allows us to contemplate reforms in the government role in the delivery of health care. In the next chapter, we examine the two largest public-sector interventions in health insurance markets—the Medicaid and Medicare programs—and the implications of past evidence for future directions in health care reform.

DO NOT COPY
reibaker@live.unc.edu

15.1 An Overview of Health Care in the United States

In 2019, the United States spent \$3.8 trillion on health care, or 17.7% of GDP.¹⁰ As noted earlier, this represents a dramatic increase from 70 years ago. This amount is also much higher than the amount spent in other industrialized nations. As [Figure 15-1](#) shows, in 2019, the United States devoted nearly twice as large a share of our economy to health care as did Finland or the United Kingdom.¹¹ Health spending in the United States amounts to \$11,582 on average for each person.¹²

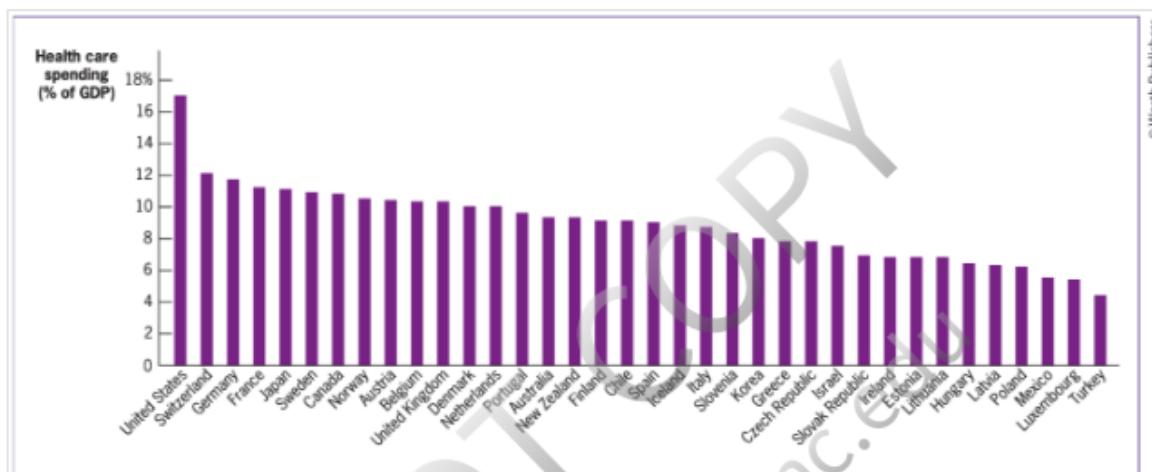
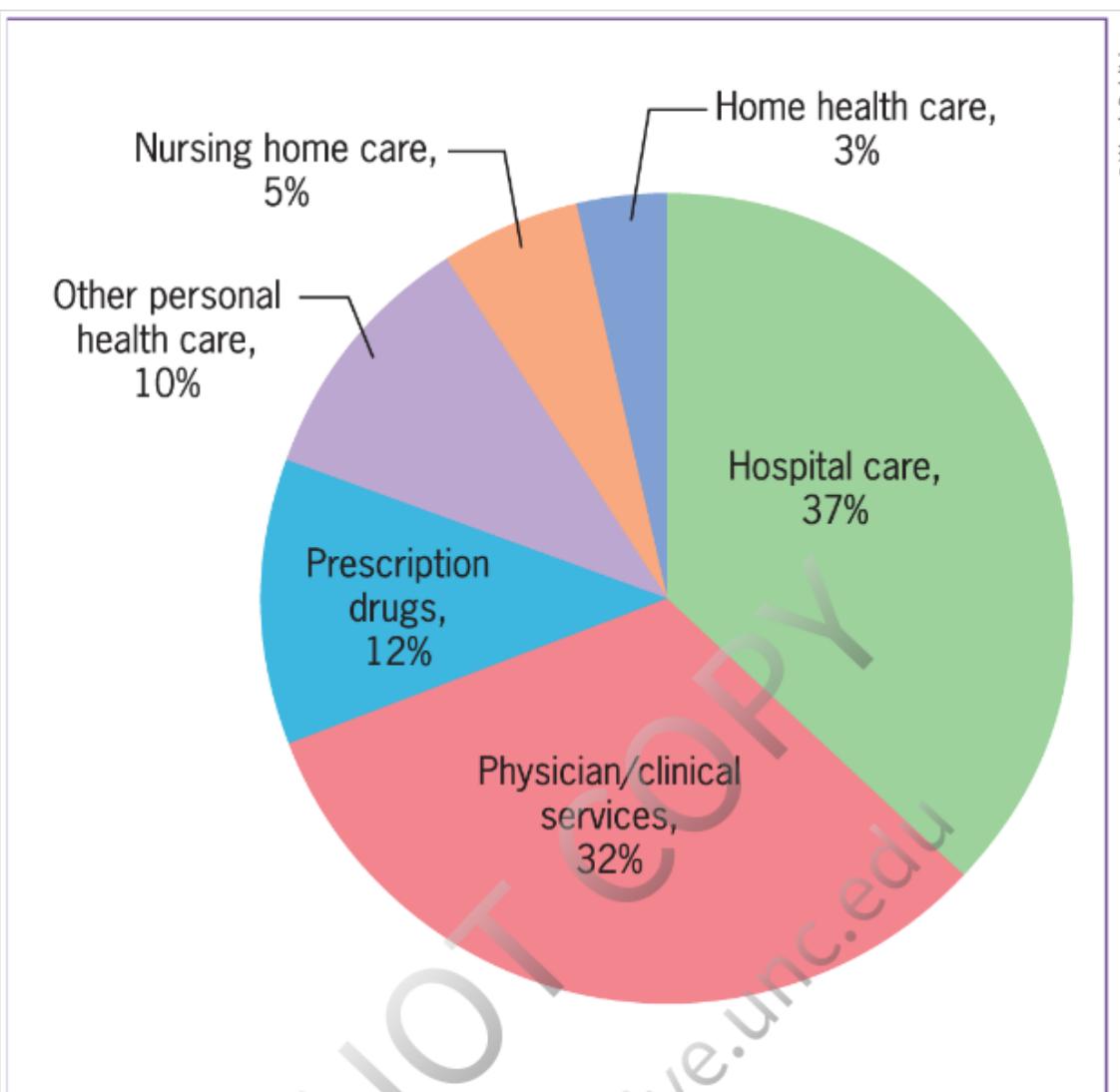


FIGURE 15-1 Health Care Spending in OECD Nations in 2019 • Health care spending is much higher in the United States than in the typical industrialized nation.

Data from: [Organization for Economic Cooperation and Development \(2021\)](#).



Where do our health dollars go? [Figure 15-2](#) shows the distribution of health spending across the major categories of expenditures. Almost 40% of the typical health dollar is spent on hospital care, and nearly a third is spent on physician care. Prescription drug spending accounts for more than a tenth of health spending, while the sum of spending on nursing home care and care for older adults in their own homes accounts for almost another tenth.



■ **FIGURE 15-2 Distribution of National Health Expenditures, 2019** • The largest single category of health care spending in the United States in 2019 was hospital care. Together hospital and physician spending accounted for over two-thirds of all health care spending.

Data from: [Centers for Medicare and Medicaid Services \(2020\)](#), Table 2.



APPLICATION

Finding the Inefficiency in U.S. Health Care



As we see in [Figure 15-1](#), the United States is a major outlier in international terms when it comes to health care spending. This is illustrated from a different point of view in [Figure 15-3](#), compiled using data from the Commonwealth Fund.¹³ This figure shows the ranking of the United States relative to other developed nations along a number of indicators of health system outcomes. There are five outcomes considered, and in each case, the U.S. value is normalized to be 1 so that other countries can be shown relative to that level. For example, the first set of bars shows that the United States has

the highest per-person health care costs of this set of countries and that countries such as the United Kingdom spend less than 45% as much per capita than the United States does. The United States also has the highest rate of infant mortality, the highest rate of preventable death, and the highest rate of going without care over the past year because of cost—only Canada has a higher wait time for physician care.

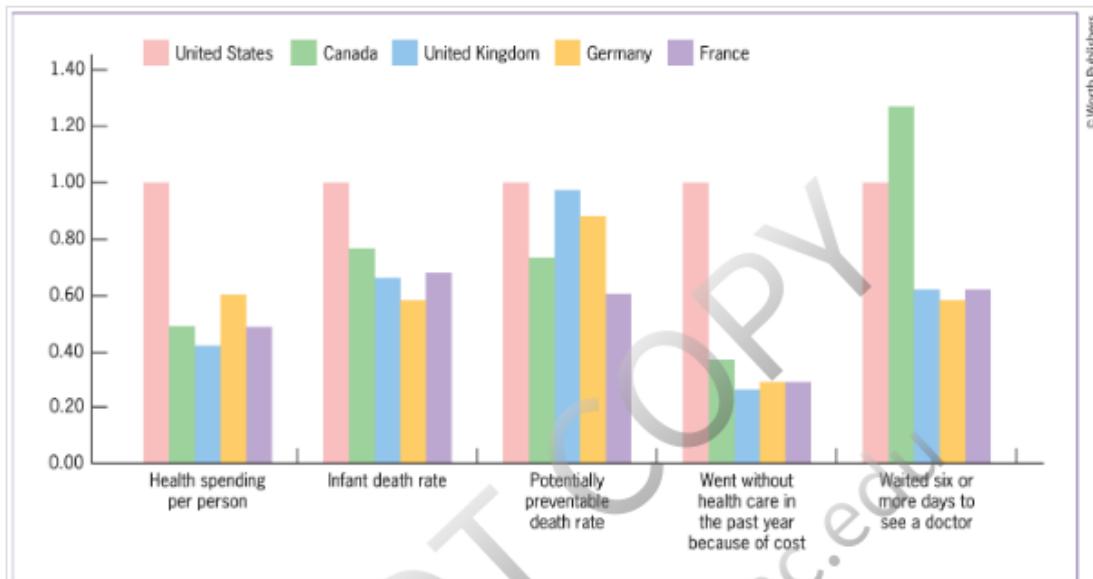


FIGURE 15-3 U.S. Rankings in Health System Outcomes • This figure shows the ranking of the United States relative to other developed nations along a number of indicators of health system outcomes. The United States spends twice as much as most other developed countries on health care, yet it is outperformed in almost every other indicator.

Data from: The Commonwealth Fund at www.commonwealthfund.org.

i

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
reibaker@live.unc.edu

In this case, if Emma values health insurance at above its cost (if there is a lot of weight on the second term in her utility function relative to the first), she might stay at her old job, even though $MP_2 > MP_1$, because of her disutility of losing insurance. Thus, health insurance availability may inhibit productivity-increasing job switches.³⁸ In fact, as we review in the Empirical Evidence box, it appears that job lock is an important phenomenon in the United States: workers with health insurance are about 25% less likely to change jobs because of that insurance.

EMPIRICAL EVIDENCE

Health Insurance and Mobility

Is job lock an important problem in reality? A large literature has investigated this question and concluded that it is. Initially, this literature compared the mobility rate of those with and without health insurance and showed that those who have health insurance are less likely to leave their jobs than those who do not, suggesting job lock.

However, these groups do not form sensible treatments and controls because they are likely to be dissimilar in at least two ways. First, those who choose to enter jobs that offer health insurance may be quite different from those who do not. For example, they may be in worse health. If worse health is associated with less job mobility, then this may be the reason for the observed correlation of health insurance and mobility: health care needs rather than insurance coverage explains why these individuals are less likely to leave their jobs. Second, jobs that provide health insurance also typically offer higher wages and other benefits such as pension plans or vacation, while lower-paying jobs may not offer health insurance. Individuals may be reluctant to leave these jobs not because they fear losing health insurance coverage but because these jobs are simply too good to leave! As a result of this lack of comparability between treatment groups (those with health insurance) and control groups (those without), the estimates are biased.

A more sophisticated literature in the 1990s surmounted this problem in two different ways.³⁷ First, studies used a difference-in-difference strategy that compared a treatment group of those who valued health insurance particularly highly with a control group of those who did not. These studies asked, for example: Does having health insurance lower the mobility rate among those who don't have any other source of insurance coverage (treatments), relative to those who do have coverage from their spouses or some other source (controls)? If job lock is an important problem, it should be found most prominently among those who don't have coverage from a spouse. Other reasons for the correlation of insurance with mobility (bias that does not represent true health insurance effects) are captured by the control group of those who do not have spousal coverage.

Second, studies examined the impact of state laws that allowed workers to continue to purchase their employer-provided health insurance for some period of time after leaving their jobs. These laws mitigated the problem of job lock to some extent because workers could be sure to have coverage for a period of time even if they left a job with health insurance for one without health insurance.

These laws were passed in some U.S. states in the 1970s and 1980s so that a quasi-experimental analysis was possible: individuals in states passing laws were the treatment group (because job lock

should be loosened), those in states without laws were the control group, and any difference in mobility was due to a loosening of job lock through these laws. Federal legislation in 1986, part of the Consolidated Omnibus Reconciliation Act, or COBRA, then made continuation coverage available nationally (which is why it is often known as COBRA coverage). The passage of COBRA provided another opportunity for quasi-experimental analysis in which those workers in states that did not already have laws were the treatment group and those in states that already had laws (and were thus unaffected by the federal law) were the control group.

The results from these studies support the notion that job lock is quantitatively important. [Madrian's \(1994\)](#) estimates, for example, suggested that it reduces mobility across jobs among those with health insurance by as much as 25%. Subsequent studies in this same vein have found that a lack of health insurance coverage for retirees reduces the odds that someone will retire before age 65 from their job because older persons do not want to risk being uninsured before they become entitled to Medicare at age 65. One major goal of the Affordable Care Act was to reduce job lock by providing nondiscriminatory coverage for those without employer insurance through the insurance exchanges described in [Chapter 16](#).

15.2 How Generous Should Insurance Be to Patients?

In considering government intervention in health insurance markets, the first question is the following: How generous should health insurance be? As with other insurance discussed in [Chapters 12, 13, and 14](#), the optimal generosity of health insurance will be determined by trading off the consumption-smoothing benefits and moral hazard costs of insurance. Yet generosity is measured in a very different way with health insurance than with the other programs we have studied. For Social Security or unemployment insurance, generosity reflects the share of pre-event wages replaced, or perhaps the duration of benefits. In the context of health insurance, generosity reflects the share of medical spending that will be reimbursed by the health insurer.



"On a scale of 1 to 10, how good is your insurance?"

"On a scale of 1 to 10, how good is your insurance?"

The generosity of health insurance is therefore measured along two dimensions. The first is generosity to *patients*: What share of the bill for medical services should be paid by the insurer, and what share should be paid by the patient through deductibles, copayments, and coinsurance? The most generous health insurance plan is one that provides [**first-dollar coverage**](#), reimbursing providers fully with no cost to the patients themselves. Plans can be less generous to consumers either by refusing to reimburse some services, so that patients pay the full cost, or by raising the amount that patients need to pay when they get the service. So, in this section, we discuss the following question: What share of a patient's medical spending should be reimbursed by the insurer, and what share should be paid by the patients themselves?

first-dollar coverage

Insurance plans that cover all medical spending, with little or no patient payment.

The second dimension of insurance generosity is generosity to *providers*: How should insurers reimburse providers for the services they deliver? Should insurers just pay

the amount billed by the provider for medical services, or should the insurer limit in some way how much the provider will be reimbursed? In the next section, we discuss this second dimension of generosity.

Consumption-Smoothing Benefits of Health Insurance for Patients

Applying what we learned in [Chapter 12](#), the benefits of health insurance to individuals are clear. Risk-averse individuals will value health insurance as a means of smoothing their consumption with respect to the cost of medical events. Not all types of medical events are created equal, however. Some are minor and predictable, such as a quick physician visit for a checkup. Others are more extensive and unpredictable, such as hospitalization for a heart attack. The key insight of expected utility theory is that insurance is much more valuable for the latter types of medical events and that there is relatively little consumption-smoothing benefit from covering the former type of (minor) events. Thus, first-dollar coverage does not provide much more consumption smoothing than does health insurance that makes patients pay the minor costs of medical care and has insurance pay only the higher costs of major medical events.

The consumption-smoothing benefit from first-dollar coverage of minor and predictable medical events is small for two reasons. First, risk-averse individuals gain little utility from insuring a small risk: the disutility to the individuals from paying insurance premiums for small risks is roughly the same as the utility they gain from insuring those risks. We can illustrate this point by returning to the initial example from [Table 12-1 \(page 340\)](#). In that example, Mimi faced an actuarially fair premium for insurance against a catastrophic risk: the 1% risk of being in a car accident and having medical expenses that wiped out her income (\$30,000). We showed that Mimi could greatly increase her utility by buying full insurance against that small but catastrophic risk.

Consider the same example now, but imagine that the medical care costs of the accident were only \$100 rather than \$30,000. In that case, Mimi's utility with no insurance is:

$$\begin{aligned} & (0.99 \times \sqrt{30,000}) + (0.01 \times \sqrt{29,900}) \\ & = (0.99 \times 173.2) + (0.01 \times 172.9) = 173.2 \end{aligned}$$

Suppose instead that Mimi buys insurance that pays the \$100 if she is hit. Because the odds of an accident are 1%, the actuarially fair premium for such insurance is \$1, which Mimi pays regardless of whether she gets hit. With insurance, Mimi's expected utility becomes:

$$\begin{aligned} & (0.99 \times \sqrt{30,000 - 1}) + (0.01 \times \sqrt{30,000 - 1}) \\ &= (0.99 \times 173.2) + (0.01 \times 173.2) = 173.2 \end{aligned}$$

Thus, *Mimi's utility does not measurably increase if she buys insurance*. This outcome stands in contrast to the earlier example, when there was a large increase in utility from buying insurance, and illustrates that insurance has little value to individuals for very small risks. Technically, the consumption-smoothing gains from insuring small risks is small because there is extremely little diminishing marginal utility for small changes in consumption: the losses from reducing consumption by a dollar are roughly equal to the gains from increasing consumption by a dollar. For risks that involve very small income loss, then, individuals are no longer particularly averse to risk (they are approximately *risk neutral*).

The second reason that the consumption-smoothing benefits of first-dollar coverage are small when medical spending is small and predictable is that individuals are much more able to self-insure such spending than to self-insure large and unpredictable medical events. Individuals can save in advance for their expected physician visit, but it would be very inefficient for them to save an extra \$200,000 against the small chance that they might have a heart attack.

Moral Hazard Costs of Health Insurance for Patients

Offsetting the consumption-smoothing benefits of health insurance to individuals is the risk of moral hazard. The classic analysis of patient-side moral hazard in health insurance is provided in [Feldstein \(1973\)](#), and illustrated in [Figure 15-7](#), in the example of a doctor's office visit (although the theory applies generally to most medical goods and services). On the horizontal axis is the number of office visits by Marty; on the vertical axis is the price that Marty must pay for each office visit. Assume that the marginal cost of producing an office visit (physician's time, supplies, etc.) is constant at \$100, so that the supply curve for medical care is the horizontal line *S*. Assume also that Marty has a downward-sloping demand curve for

health care such as office visits: she is willing to pay less for more health care (due to diminishing marginal utility). Finally, assume that Marty is in a large group, so that her medical spending is irrelevant for the insurance premium that she pays. (There is no individual-level experience rating in insurance prices.) The only costs to her for using medical care, then, are the copayments that she makes for that care.

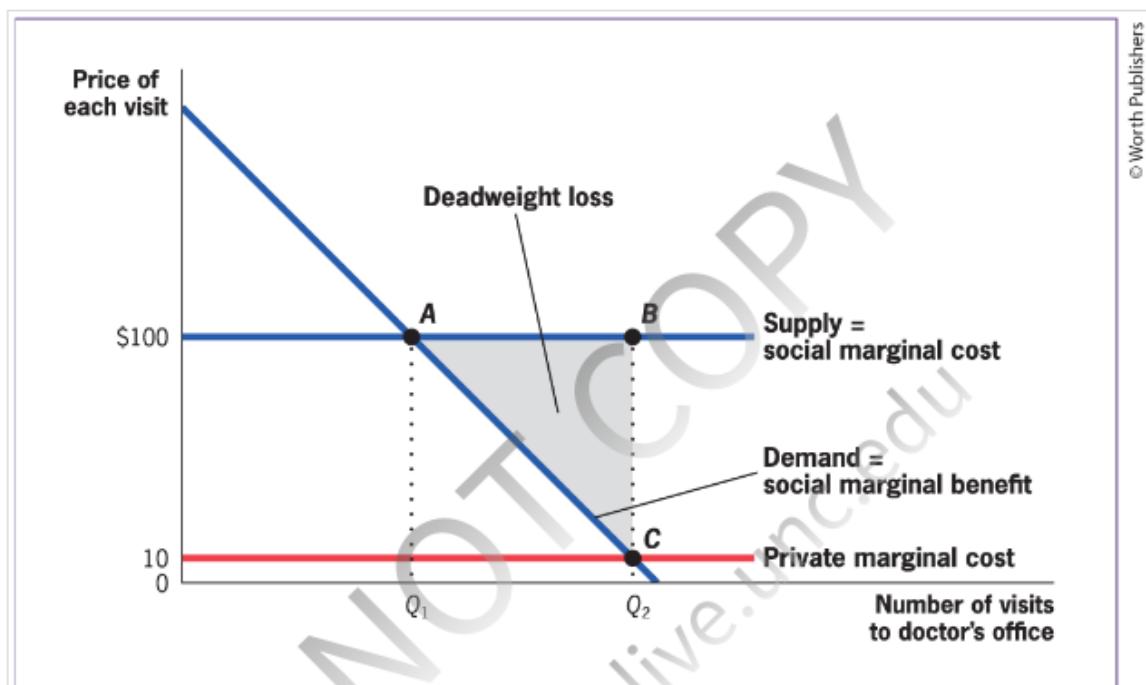


FIGURE 15-7 Patient-Side Moral Hazard • With no insurance, and a cost of \$100 per visit, individuals would consume Q_1 doctor's office visits, where marginal costs and benefits are equal. With only a \$10 copayment, however, individuals consume Q_2 worth of visits, where private marginal costs equal social marginal benefit; this overconsumption of health care leads to a deadweight loss of ABC.



If Marty faced the full price for medical care (i.e., if she had a 100% coinsurance rate), she would consume Q_1 visits per year (point A). At point A, her marginal benefit of receiving care (summarized by the demand curve) equals her full price of that care, \$100. Q_1 is also the socially optimal level of medical care; at this point, social marginal benefits (demand) equal social marginal costs (supply).

Now suppose that Marty must pay a \$10 copayment when she goes to the doctor, as is typical in many health plans today. In this case, she will face a private marginal cost of only \$10 and will choose Q_2 visits (point C). Q_2 is an inefficiently large amount of medical care because at this quantity, private marginal costs (\$10) are far below social marginal costs (\$100). Thus, there is an inefficiency of the area ABC, which represents all of the units of medical care that are delivered with a marginal benefit (demand) below their social marginal cost. The moral hazard associated

benefit (demand) below their social marginal cost. The moral hazard associated with not charging Marty for the full cost of care leads to a deadweight loss of the area ABC .

The benefit of this small copayment is consumption smoothing: individuals such as Marty value the fact that they are insured against paying the full cost when they go to the doctor. A 100% coinsurance rate may induce efficient medical care use, but it provides no insurance against large medical costs, which is valued by risk-averse consumers. Here we see the fundamental trade-off of health insurance: the gains in terms of consumption smoothing (paying \$10 instead of \$100 when you go to the doctor) versus the costs in terms of overuse of medical care (consuming Q_2 instead of the socially optimal Q_1).

The “Flat of the Curve”

This inefficient overuse of medical care has led some to claim that we practice “flat of the curve” medicine. This notion is illustrated in [Figure 15-8](#), which graphs the relationship between medical spending and the associated improvement in health, or the “health effectiveness curve.” The horizontal axis in this figure measures the *level* of medical spending. The vertical axis measures the *marginal health benefit* from the next dollar of medical spending. Health benefits in this stylized figure are represented in dollar terms, or the monetary value of improving one’s health; an improvement of \$1 means that this improved health is worth \$1 to the individual. Each point on this curve is the marginal improvement in health for spending the next dollar on health care.

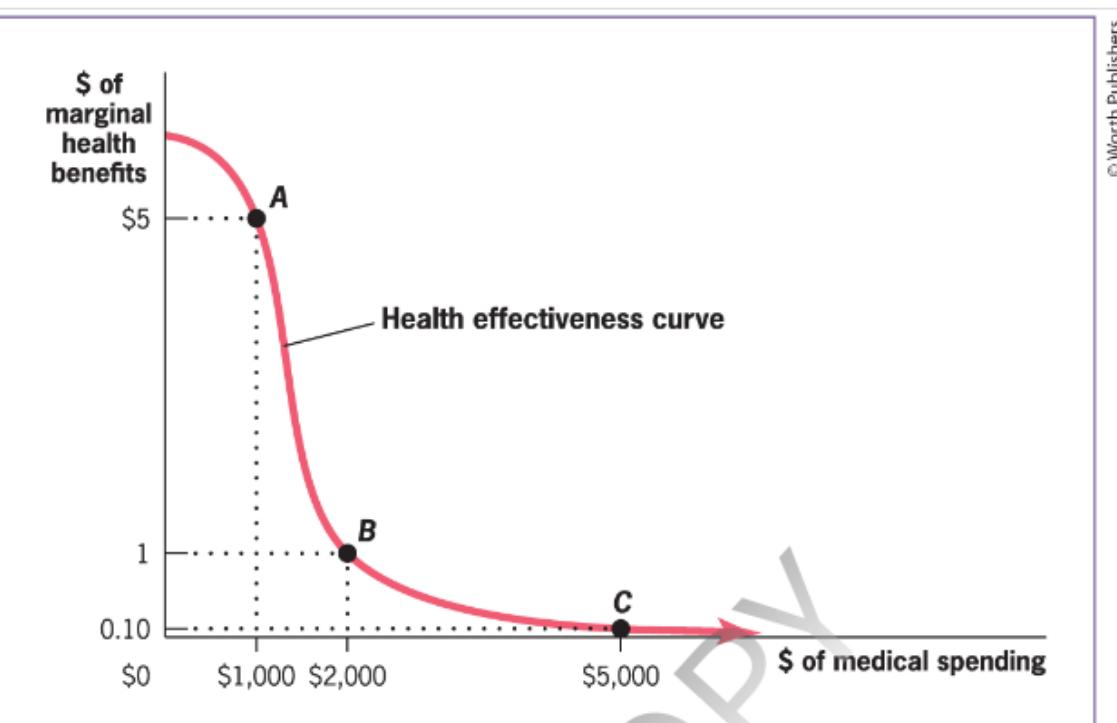


FIGURE 15-8 The “Flat of the Curve” • Spending on health care is assumed to initially be very productive in terms of improved health care outcomes, but that productivity dwindles as spending rises. The curve shows the value of improved health for each dollar in medical spending. At point A, when individuals are spending \$1,000 on health care, each dollar of medical spending buys \$5 worth of improved health; at point B, when individuals are spending \$2,000 on health care, each dollar of medical spending buys \$1 worth of improved health. Beyond point B, however, there is much less than \$1 in improved health for each \$1 in medical spending.

i

Initially, health spending is very productive in terms of improving health status because there are a series of very cost-effective medical interventions (such as vaccination against influenza for older adults). Point A, for example, measures the marginal health benefits to spending another dollar of health care once one has spent \$1,000 on health care. This next dollar of spending improves the individual’s health at that point by \$5 (vertical axis), or five times the increase in health spending. As health spending rises, however, we move from clearly cost-effective interventions to less clearly cost-effective interventions. At point B, moving from \$2,000 to \$2,001 of health care spending improves the value of health by \$1, equal to the spending increase. At point C, moving from \$5,000 to \$5,001 in health care spending improves health by only \$0.10, or one-tenth the spending increase. Eventually, additional spending does no good in terms of improving health, and the effectiveness curve flattens out; some claim that overuse may actually reduce one’s health, with the curve dropping below zero at very high levels of spending.

Optimally, people should stop getting medical care when the additional health benefit is smaller than the additional medical cost. When we look at [Figure 15-8](#), we see that people should not get medical care beyond point *B*, the point at which each dollar of spending buys a dollar of improved health. If individuals paid the full cost of their health care, point *B* would be the socially optimum level of health care that would be chosen in a competitive market. If individuals do not pay much for their additional health care, however, they will demand health care as long as the effectiveness curve is not perfectly flat. This demand pushes our society to the right of point *B*, into the region where each dollar of medical care buys much less than \$1 in improved health. As a result, some studies have judged that as many as one-third of all medical procedures delivered in the United States are “of questionable benefit.”³⁹

How Elastic Is the Demand for Medical Care? The RAND Health Insurance Experiment

The extent to which moral hazard causes the actual quantity of health care consumed to exceed the socially optimal quantity depends on both the copayment amount and the elasticity of demand for medical care. Many years ago, policy makers assumed that this elasticity was close to zero; individuals went to the doctor when they were sick and didn’t if they weren’t sick. Several decades of empirical economics research has shown this not to be the case.

The best evidence on the elasticity of demand for medical care comes from one of the most ambitious social experiments in U.S. history: the RAND Health Insurance Experiment (HIE), which was conducted in the mid-1970s at several sites in the United States. In the HIE, individuals were randomly assigned to plans with different coinsurance rates. Some participants were placed in plans with no coinsurance (basically, a free care plan), others were assigned to plans with a coinsurance rate of 95% (these individuals paid 95% of all their health costs), and the rest were assigned plans with a range of coinsurance rates in between. Because individuals who were assigned less-generous plans had to pay more (which is unethical and might have made it hard to recruit participants), each plan had an “out-of-pocket maximum” of \$1,000. Once individuals had spent \$1,000 on their medical care, they did not have to pay any additional costs of care, regardless of their plan.

EMPIRICAL EVIDENCE

Estimating the Elasticity of Demand for Medical Care

Initial research on the elasticity of demand for medical care proceeded by comparing individuals in different types of health plans. Some health plans provided first-dollar coverage, where individuals had no coinsurance; others had large coinsurance, with coinsurance rates of 20% (the patient pays one-fifth of the cost of the visit) or more. These studies showed that individuals in plans with higher coinsurance used less care, which suggested that medical care demand was somewhat elastic.

It is likely, however, that these earlier studies were seriously biased because insurance plans are not randomly assigned to individuals. Individuals who spend a lot on health care are likely to buy plans with low coinsurance, so the correlation between low coinsurance and high use of medical care is not a causal effect of coinsurance on utilization. Instead, the correlation reflects that high-utilization individuals have chosen low-coinsurance plans. This is a classic example of the bias problem that arises when trying to assign causal interpretations to correlations.

The RAND HIE was designed to address this shortcoming. Because of random assignment, we can use the data from the HIE to assess the causal impact of coinsurance on utilization by comparing individuals in different plans. By definition, individuals with different coinsurance rates were identical other than for their coinsurance rates, so it is possible to compare the utilization of those with high-coinsurance rates and low-coinsurance rates to identify the elasticity of demand for medical services.⁴⁰

Unfortunately, we have been unable to run another randomized trial of insurance generosity in the 40 years since the RAND HIE.⁴¹ But recent studies have used quasi-experimental approaches to estimate the price elasticity of medical demand. These studies compare utilization before and after a copayment change in a plan, relative to other plans that do not change copayments. These studies have the weakness that the copayment changes are not experimental; for example, firms may raise copayments for prescription drugs in the face of rising demand, leading to the appearance that higher copayments lead to more drug utilization! However, a large number of studies in many different settings have now been completed of this type, and the results are fairly uniform. These studies confirm the RAND HIE conclusion that higher copayments reduce care, particularly for prescription drugs. They also find that the reduction in care has deleterious effects on health for the chronically ill, highlighting the value of targeted copayment policies.

A quasi-experimental study in Japan confirms that these types of results hold in other nations as well. As [Shigeoka \(2014\)](#) described, the structure of cost sharing for insurance in Japan is such that cost sharing drops precipitously at age 70. Before that age, patients face a 30% coinsurance rate, with an out-of-pocket maximum exposure of more than 80,000 yen; once they reach age 70, however, the coinsurance drops to 10% with a maximum out-of-pocket exposure of only 12,000 yen. Thus, there is a sharp reduction in the burden faced by patients receiving care at age 70.

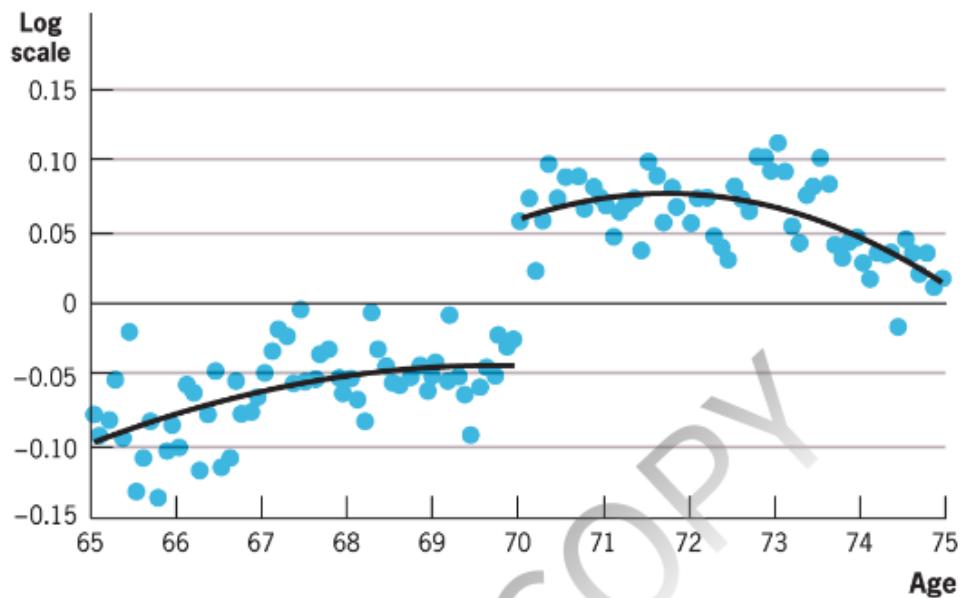
There is, in fact, a corresponding increase in health care spending at age 70 that goes much beyond typical aging patterns. This is illustrated in [Figure 15-9](#) for both outpatient visits (physicians and nonadmitted patient hospital spending), shown in panel (a), and inpatient hospital admissions, shown in panel (b). Each dot in this figure represents the actual average of the natural logarithm of the number of visits to admissions by month of age (e.g., the first dot is 65 years old exactly, the second 65 years and one month, and so on).⁴² In both cases, the number of outpatient and inpatient visits increases with age, which is consistent with a growing need for medical care as we get older. But there is an enormous jump at age 70, exactly when cost sharing drops. The fitted lines in the figure show the magnitude of this “regression discontinuity,” as discussed in the Empirical Evidence box on “The Effect

© 2014 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has determined that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

of Legal Drinking at Age 21" in [Chapter 6 \(p. 170\)](#). Estimates in Shigeoka's study showed that there is a 10% jump in outpatient visits at age 70 and an 8% increase in inpatient admissions.

© Worth Publishers

(a) Age profile for outpatient visits



(b) Age profile for inpatient admissions

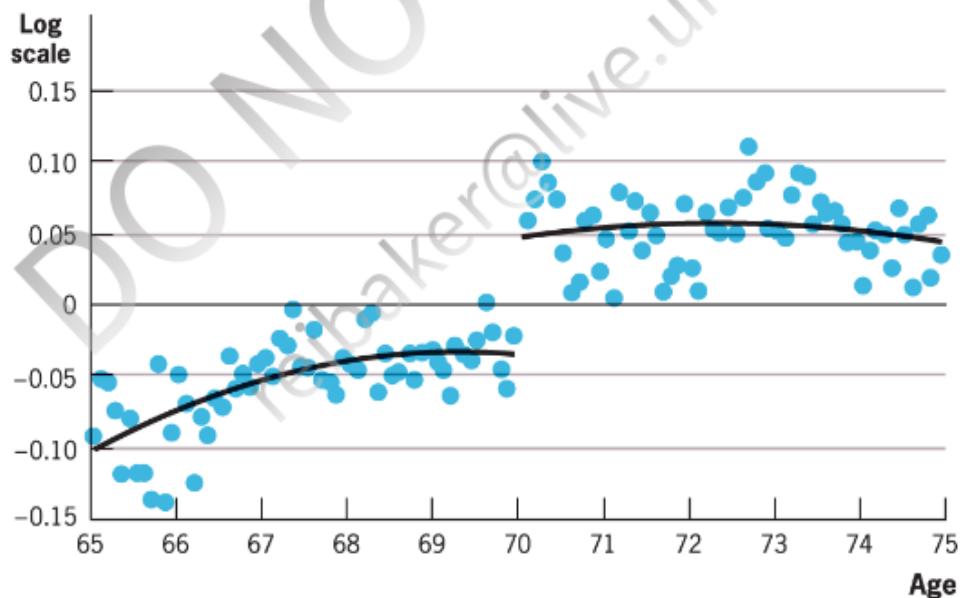


FIGURE 15-9 Age Profiles for Outpatient Visits and Inpatient Admissions Panels (Log Scale) • (a) (outpatient visits) and (b) (inpatient hospital admissions) show that both increase with age, which is consistent with a growing need for medical care as we get older. Age 70 marks a huge drop in cost sharing for the elderly and a corresponding jump in visits to the doctor and hospital admissions.

Data from: Shigeoka (2014), Figures 2A and 4A.

Data from: [Shigeoka \(2014\)](#), Figures 2A and 4A.



Strikingly, the implied elasticity from these changes is almost exactly what was found in the RAND HIE: an elasticity of roughly -0.2 . Moreover, there is no measurable effect on patient mortality; mortality shows no discontinuity at age 70 despite the very large increase in medical care utilization. This confirms the “flat of the curve” conclusion from the HIE.

The findings of the HIE were striking.⁴³ First, medical care demand is price sensitive: individuals who were in the free care plan used about one-third more care than those paying 95% of their medical costs. The implied elasticity across the entire study was 0.2, meaning that each 10% rise in the price of medical care to individuals led them to use 2% less care.

Second, those who used more health care due to the lower price did not, on average, see a significant improvement in their health. This finding suggests that the typical person is indeed on the flat of the health effectiveness curve when responding to changes in coinsurance. The finding *does not imply*, however, that insurance isn’t valuable at all because everyone in this experiment was insured; after a family’s health spending reached \$1,000, it had full insurance. The RAND results imply that after individuals are insured for large expenditures, varying the coinsurance for small expenditures does not seem to affect their health on average.

Third, for those who are chronically ill and don’t have sufficient income to easily cover copayments, there was some deterioration in health. In particular, low-income individuals who were hypertensive (had high blood pressure) saw dangerous increases in their blood pressure arising from lack of care. More recent studies have confirmed that finding, concluding that those who have treatable but chronic illnesses may be made worse off by copayments. Indeed, some studies find that for the chronically ill, raising copayments actually raises total medical costs because the reduced use of prescription drugs and physician visits results in more expensive hospitalizations due to health deterioration.⁴⁴

Optimal Health Insurance

The finding of significant deadweight loss from moral hazard in the health insurance market suggests that the optimal health insurance policy is one in which individuals bear a large share of medical costs within some affordable range and are only fully insured when costs become unaffordable. This structure is optimal because first-dollar coverage has little consumption-smoothing benefit but a large

moral hazard cost.

As we showed, the consumption-smoothing benefit of first-dollar coverage is small. But first-dollar coverage also has substantial moral hazard cost because it encourages individuals to overuse the medical system, demanding care for which the social costs exceed the social benefits. In other words, coverage for small amounts of medical spending has little benefit (because individuals don't much value consumption smoothing for small risks) and significant cost (because individuals are using care where marginal benefit is less than marginal cost), so the optimal insurance plan should not provide such coverage. Rather, it should insure only large medical expenses for which the consumption-smoothing gains are large (and for which the moral hazard might be smaller, such as for heart attacks). An example of an optimal insurance plan is [Feldstein's \(1973\)](#) "Major Risk Insurance" plan, in which individuals would make a 50% copayment on all services until they spent 10% of their income on medical care, beyond which there would be no more copayments.

Results from the RAND HIE and subsequent studies, however, suggest an important caveat to that general structure: there is a benefit to targeting consumer cost sharing to encourage appropriate care and discourage inappropriate care. For example, reduced cost sharing can be cost effective for prescription drugs for those with chronic illness. Indeed, it may even be optimal to subsidize preventive care for that population in order to avoid more expensive care down the road. There is a growing interest in such "value-based insurance design" as a means of incorporating consumer decision making into improving health care; a recent review of dozens of studies found that such a design can increase medication adherence without an increase in total health care spending.⁴⁵

Countervailing Pressure: the Tax Subsidy to ESI

While optimal health insurance may feature high cost sharing, at least on some services, employer-sponsored insurance in the United States has traditionally been much more generous. One reason is that health insurance expenditures by employers are tax subsidized; that is, payments to employees in the form of wages are taxed, while payments in the form of health insurance are not. This is a major reason health insurance is offered through firms. It also has the effect of increasing consumption of health insurance relative to other goods by lowering the relative price of health insurance. For the typical worker, the marginal dollar spent on better health insurance buys a dollar's worth of insurance, while the marginal dollar paid in wages buys only \$0.70 of other consumption because \$0.30 of that dollar goes to

pay taxes. This situation will lead employees to devote a larger share of their compensation to health insurance than they would in the absence of the subsidy; even if individuals would not demand generous insurance at unsubsidized prices, they may choose it when there is a tax subsidy for its purchase. Estimates suggest that eliminating the tax subsidy to health insurance purchases by employers could lower employer-provided insurance spending by almost half.⁴⁶ [Powell \(2016\)](#) estimates that the distortion to health insurance plan design from the tax subsidy (which induces plans to excessively cover low medical expenditures) leads to a deadweight loss of \$34–\$44 billion per year.

This estimate does not imply that removing the tax subsidy today would necessarily be good health policy. Many individuals (more than 20 million by one estimate) would lose their employer-provided health coverage if the subsidy were eliminated and would then face the difficulties of obtaining insurance in the nongroup market, with its high prices and incomplete coverage.

One compromise policy that has been suggested is to *cap* the tax subsidy. Ending the tax subsidy would be accomplished by taxing workers on all employer contributions to health insurance as if these were earnings to the workers. Capping the tax subsidy would mean continuing to exclude from income taxation employer health insurance spending up to some level (perhaps the median cost of a health insurance plan in the area) but taxing as worker earnings any employer spending above that level. For example, your employer would add to your reported wage income any amount spent on insurance above the ceiling level, and you would be taxed on this in the same way you are taxed on wages. This approach would retain the tax subsidy to employers for some basic insurance spending, thus encouraging them to provide some insurance coverage. The subsidization to excessively generous coverage, however, would end. There would no longer be a reason for employees to choose a very generous insurance plan over cash compensation, once employers are spending the cap amount. The “Cadillac Tax” imposed by the ACA has similar incentives to such a cap, as discussed in [Chapter 16](#).

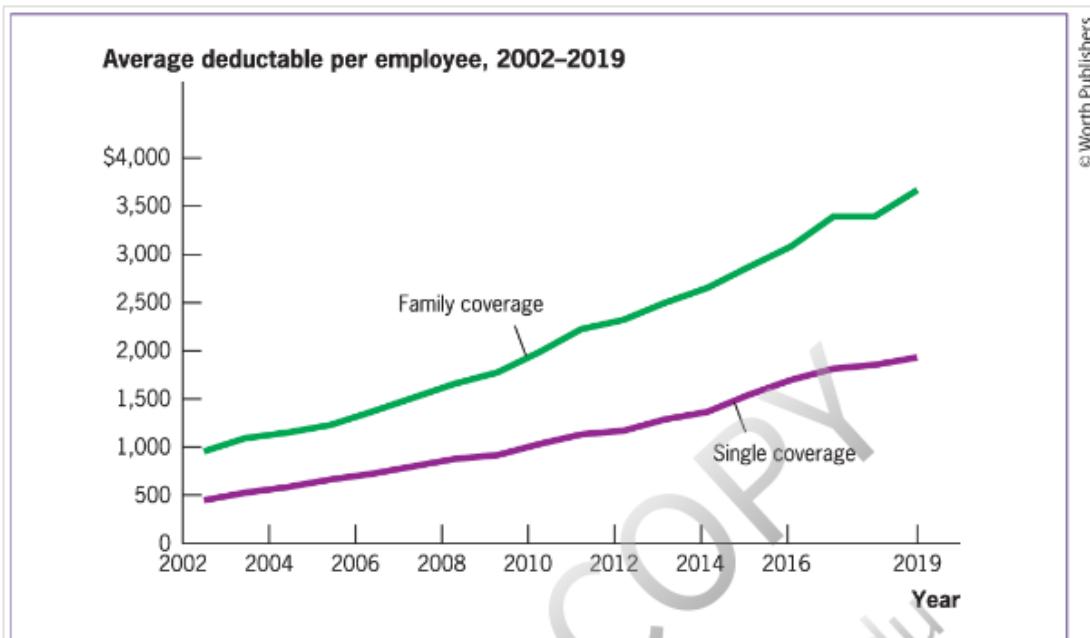
APPLICATION

The Growth of High-Deductible Plans in the United States



For many decades in the United States, most private insurance was not structured very similarly to the optimal health insurance principles laid out above. [Figure 15-10](#) shows the average deductible in

the optimal health insurance principles laid out above. [Figure 15-10](#) shows the average deductible in employer-sponsored health insurance plans in the United States. Deductibles, and other forms of cost sharing, were very low through the early 2000s, and “first dollar” insurance coverage was the norm. That has changed dramatically over the past two decades. Today, the average employer plan has a deductible of \$1,930, and 25% of plans have deductibles of at least \$2,700 per year.



■ **FIGURE 15-10 Average Deductibles for Employer-Sponsored Health Insurance** • The average health insurance deductible has risen steadily from 2002 to now.

Data from: [Agency for Healthcare Research and Quality \(2021\)](#).



Part of the reason for the growth in high-deductible plans can be traced to the use of [Health Savings Accounts \(HSAs\)](#). Introduced as part of the Medicare prescription drug legislation of 2003 (discussed in detail in the next chapter), HSAs are savings accounts tied explicitly to health insurance plans with high deductibles (at least \$1,000 for individuals and \$2,000 for families per year). Money in HSAs can be withdrawn to pay out-of-pocket medical costs, including deductibles. Importantly, the amount put into the HSA is not taxed—so spending on out-of-pocket costs becomes treated in the same way for tax purposes as does spending on insurance premiums (as opposed to other insurance plans, where insurance spending is tax subsidized, but out-of-pocket spending is not).⁴⁷ Individuals who do not spend down their accounts in any year can roll the money over to the next year, and any money left at retirement can be withdrawn either for the retiree’s health care needs, in which case, it’s tax-free, or for any other purpose, in which case it is taxed. Thus, HSAs provide a large tax subsidy to those whose plans have a high deductible, hopefully encouraging increased use of such insurance arrangements. HSAs are rapidly growing in popularity, covering more than 29 million Americans in 2020.⁴⁸

health savings account (HSA)

A type of insurance arrangement whereby patients face large deductibles, and they put money aside on a tax-free basis to prepay these deductibles.

A host of recent studies, however, suggest potential problems with the growth in high-deductible plans. [Brot-Goldberg et al. \(2017\)](#) study the impact on employees at a large company that required all

of its employees to switch to a high-deductible plan. They find that this switch caused a substantial reduction in total firm-wide health spending (relative to control firms with no change in the plan), with an implied elasticity similar to what was found in the RAND HIE. But they have two worrisome findings. First, they find no evidence that enrollees were induced to price shop across services, one of the major arguments proposed by advocates of high-deductible plans. Second, they find that consumers reduce their use of medical care across the spectrum of health care services, including potentially high value care such as preventive services.

[Wharam et al. \(2019\)](#) provide particularly worrisome evidence of the impact of high-deductible plans in deterring needed preventive care. They find that women whose employers switched from low- to high-deductible plans saw significant delays in their time to breast imaging, biopsy, breast cancer diagnosis, and chemotherapy. They did find somewhat smaller, although still significant, delays for higher than for lower income women.

Of course, “high deductible” is arbitrary—even small amounts of cost sharing can matter when enrollees are low income. [Gross, Layton, and Prinz \(2020\)](#) study what happens to low-income Medicare enrollees who face small copayments on their drugs. They find that the arrival of Social Security checks significantly increases fills of important prescription drugs through providing “liquidity” to low-income consumers who find even these small copayments to be burdensome.

So high-deductible plans pose a trade-off. On the one hand, they significantly lower medical spending—indeed, [Rabideau et al. \(2021\)](#) find that they reduce spending on “low-value services” such as imaging at a much higher rate than they reduce other spending. On the other hand, at least some of the reduction is coming through “high value services” such as preventive care.

Fortunately, there is some evidence that the use of value-based insurance designs can potentially offset the negative effect of high-deductible plans on use of high-value care. A novel study by [Reed et al. \(2017\)](#) considered the effects on switching to a higher deductible plan among those who did and did not have an additional plan that offered free chronic disease medication. They find that among those without the additional plan, the adherence of the chronically ill with their medication regime fell significantly—but for those with the value-based plan that covered the costs of these medications, adherence did not fall. At the same time, [Gruber et al. \(2020\)](#) show that the other side of value-based incentives, charging higher deductibles specifically for low-value services such as imaging or spinal surgeries for pain significantly reduced the use of those services.

Taken all together, the evidence strongly supports the notion that high-deductible plans can reduce utilization significantly, but may do so too indiscriminately. More sophisticated pairings of high-deductible plans with value-based incentives to use high-value care and reduce low-value care may be necessary to ensure the proper balance between reducing overuse and promoting health. And incorporating income-specific aspects of cost-sharing systems, as the Affordable Care Act does with its subsidized health exchanges, may be critical to protect the lowest income enrollees. ■

The Access Motive

A second reason insurance may be so generous is that the traditional analysis overstates the costs of moral hazard. In traditional models of insurance, the additional medical care used is all due to moral hazard. [Nyman \(1999\)](#) highlights a problem with this view: some of the additional medical care used because of

problem with this view: some of the additional medical care used because of insurance is not due to moral hazard but rather to the fact that individuals can now

afford better treatments. Suppose, for example, you have an illness that costs \$1 million to treat. Suppose further that you care enough about your health that if someone handed you \$1 million today, you would immediately use the money to treat your illness. Likewise, if you had insurance that covered the treatment of illness, you would also get it treated. In this case, there is no moral hazard because all insurance has done is allow you to afford a treatment you value.

Moral hazard is technically the difference between what health care you buy with \$1 million of insurance or with \$1 million of cash. If, on the one hand, you would get treated when you have insurance, but if handed the actual cash, you would use it for other expenditures, then insurance is causing moral hazard. If, on the other hand, you would buy the expensive treatment with either the insurance or the cash, then there is no moral hazard. Insurance has allowed you to transfer sufficient resources from the situation (state) in which you are healthy to the situation that occurs when you are ill. Moral hazard is measured only by the *substitution effect* of social insurance programs, that is, the extent to which you change behavior because relative prices are changing. The *income effect* of social insurance programs, the extent to which you change behavior because you are richer, is not moral hazard.

This analysis suggests that some of the deadweight loss associated with increased expenditures due to insurance is not pure moral hazard but an income effect. As a result, we are overstating the deadweight loss associated with health insurance, which is why insurance may be more generous than our analysis would suggest is optimal.

Unfortunately, it is difficult to decompose the increase in the use of medical care into access and moral hazard effects. The access motive is a more important consideration for very expensive treatments and not for everyday treatments, such as physician's visits. Thus, the access motive may explain the strong demand for catastrophic limits on spending out of pocket, but it is unlikely to explain why individuals don't face higher copayments for minor medical procedures, such as those performed in a physician's office.

Psychological Motivations

Finally, the third reason insurance may be so generous is that there may be motivations for holding insurance that go beyond the simple expected utility model developed in [Chapter 12](#). For example, individuals with self-control problems may

developed in [Chapter 12](#). For example, individuals with self-control problems may use insurance as a commitment device. You may know that without insurance, you

will spend all of your money today and save nothing for possible future medical expenditures. By buying insurance, you are effectively forcing yourself to save for illness. You may be buying overly generous insurance because, otherwise, your impatience would leave you with no money at all to pay large medical bills that may arise. Once again, however, this is unlikely to explain the aversion of patients to modest coinsurance for minor medical events.

Alternatively, it may be that individuals simply don't like associating financial transactions with medical care. They would rather pay higher insurance up front to avoid dealing with the difficult decisions about whether they want to pay at the time of care. Clearly, more analysis is needed to understand the causes of overinsurance in America today.

DO NOT COPY
reibaker@live.unc.edu

Alternatively, it may be that individuals simply don't like associating financial transactions with medical care. They would rather pay higher insurance up front to avoid dealing with the difficult decisions about whether they want to pay at the time of care. Clearly, more analysis is needed to understand the causes of overinsurance in America today.

DO NOT COPY
reibaker@live.unc.edu

15.3 How Generous Should Insurance Be to Medical Providers?

The other type of moral hazard that is relevant for health care is on the provider side. Even if insurers could perfectly assess your true level of illness, they cannot always perfectly assess how much it costs to treat that illness. For many illnesses, there is no clearly delineated course of treatment, and major illnesses can progress very differently in different individuals. As a result, insurers have traditionally reimbursed medical providers according to their reported costs of treatment; if two physicians billed different amounts to treat a heart attack, they were typically paid those different amounts, even if characteristics of the heart attack patients were similar. Such a reimbursement system is called **retrospective reimbursement** because insurers simply reimburse physicians for the costs they have already incurred.

retrospective reimbursement

Reimbursing physicians for the costs they have already incurred.

Retrospective reimbursement removes any incentive for providers to treat their patients cost effectively. Suppose that all a doctor wants to do is make patients as healthy as possible. In that case, if there is a medical procedure or test that has any possible medical benefit that exceeds the private costs to the patient (the coinsurance plus any discomfort or time lost due to the procedure), the doctor will provide that care. After all, why skimp when you are risking someone's health? The problem is that the *social* cost of these procedures may greatly exceed the private costs (and the private benefits) to the patient, particularly if health care is being delivered on the "flat of the curve." Thus, there is overuse of medical care and a deadweight loss from provider-side moral hazard.

Moreover, the overuse problem can be exacerbated if providers care about their incomes as well. Suppose that your doctor wants to not only maximize your health but also maximize his or her income. In that case, if there is a procedure that does



you no good (but no harm or at least not much harm), and on which a net profit can be made, then your doctor will undertake that procedure.

Managed Care and Prospective Reimbursement

The twin problems of patient- and provider-side moral hazard were assumed to be major drivers of the rapid rise in health care costs during the postwar period. Initially, the insurance market responded to these moral hazard concerns through increased patient cost sharing, but this approach did not slow the rise in health care costs. Throughout the 1960s, 1970s, and 1980s, health care costs were rising sharply in real terms.

In the late 1980s and 1990s, the private market (as well as public insurance programs) turned to an alternative approach to cost control: [managed care](#), which implemented supply-side controls on the delivery of medical care. Managed care comes in two forms.

managed care

An approach to controlling medical costs using supply-side restrictions such as limited choice of medical provider.

Preferred Provider Organizations

One fundamental failure in medical markets is that it is very difficult to shop for a medical provider. How can a consumer of medical care effectively compare hospitals or physicians? And how can consumers possibly shop for cost-efficient providers even if they wanted to when prices for specific procedures are not posted by hospitals or physicians? This dilemma is particularly true for emergency care. No one will tell the ambulance driver to “take me around to a few hospitals to see which is cheapest.” This lack of price shopping means there are few competitive pressures on providers’ pricing decisions, and thus providers have no incentives to lower medical care costs.

In the 1980s, a new type of health organization called the [preferred provider organization \(PPO\)](#) gained popularity as a means of solving this problem. Essentially, PPOs serve as an intermediary that shops across providers on behalf of the insured, striking deals that can lower the cost of care. For example, a PPO might first go to an area employer and offer it 20% savings off its medical care costs if it is willing to restrict the set of hospitals its employees can use to the PPO’s hospital “network.” The PPO would then tell area hospitals that if they wanted the business

from this firm (or other firms using that PPO), they would have to cut prices by 20%. The hospitals that agreed would lower their prices for the PPO client in return for being part of the restricted PPO network. The employer would tell its employees that they could use only those hospitals in this PPO network. In principle, this shopping strategy could lead not only to lower prices but also to more efficient delivery of medical care, in the same way that shopping in other markets increases efficiency.

preferred provider organization (PPO)

A health care organization that lowers care costs by shopping for health care providers on behalf of the insured.

Health Maintenance Organizations

The other type of managed care organization is a **health maintenance organization (HMO)**. As with PPOs, HMOs restrict enrollee choice of medical providers, but HMOs go one step further by integrating the insurance and the delivery of medical care. In the classic *staff model*, HMOs hire their own physicians and may have their own hospitals. They put the providers on a salary that is independent of the amount of care they deliver. This approach removes any income incentive for the delivery of excess care.

health maintenance organization (HMO)

A health care organization that integrates insurance and delivery of care by, for example, paying its own doctors and hospitals a salary independent of the amount of care they deliver.

The more typical HMO model is the Independent Practice Association (IPA), in which the HMO contracts with independent providers (within a restricted network) to deliver care to its enrollees. In this case, HMOs counteract moral hazard through the use of **prospective reimbursement**, the practice of paying providers based on what treating patients should cost, not what the provider spends. The HMO might, for example, pay a primary-care physician (e.g., a family doctor) \$100 per month for each person in the doctor's practice, regardless of how much medical care that person uses.⁴⁹

prospective reimbursement

The practice of paying providers based on what treating patients should cost, not on what the provider spends.

Prospective reimbursement completely reverses the financial incentives of physicians. Under retrospective reimbursement, where the physician is reimbursed for all billed costs, the more care the physician delivers, the more money they

For all these costs, the more care the physician delivers, the more money they make. Now, the *less* care a physician delivers, the more money they make. The physician gets the \$100 no matter what is done for the patient, so by delivering less care, the physician can pocket a larger share of that payment. Thus, just as retrospective reimbursement offers financial incentives for excessive care, prospective reimbursement offers financial incentives for insufficient care.

Prospective reimbursement can come in many different forms, and there is indeed an enormous variety of prospective payment schemes used by HMOs. Some HMOs augment the flat payment to primary care providers with disincentives to use specialty care or hospital visits. For example, the HMO might reduce compensation every time the primary care physician recommends a specialist or a hospital visit or raise compensation if such recommendations aren't made. In this way, HMOs try to incentivize individual providers to not only limit the care delivered but also limit the care delivered by the system as a whole.

Over the late 1980s and 1990s, the vast majority of insured persons in the United States moved into some form of managed care along the spectrum from PPOs to staff-model HMOs. By 2009, 98% of the privately insured were in managed care plans.⁵⁰

The Impacts of Managed Care

The key question to consider is whether managed care has improved the functioning of the insurance market. In particular, has managed care lowered the deadweight loss of excessive medical care utilization? If so, has prospective reimbursement perhaps even gone too far and begun to restrict medical care that is valued by individuals at its cost? There is now a very large amount of literature in health economics that addresses these questions.

Spending

The consistent finding in the literature is that HMOs spend much less per enrollee than do traditional retrospective reimbursement plans. Interpreting this finding is not as straightforward as it seems, however. It is difficult to compare spending per person in these two types of plans because different types of people enroll in the plans. Many studies have found that managed care plans attract the healthiest enrollees, as measured, for example, by how much the individuals were spending before they switched to the HMO. This type of *selection* of the lowest cost individuals into HMOs means that there is not a simple comparison between the HMOs and

traditional insurance plans. Even if HMOs didn't change the delivery of medical care, their costs would be lower because they have the healthiest enrollees.

In assessing the impact of HMOs on costs, there have been a number of attempts to control for this selection problem. Perhaps the most convincing is evidence from the RAND HIE previously discussed, which had a second component that randomly assigned individuals into one of the nation's earliest HMOs, in Puget Sound, Washington. Because individuals were randomly assigned, they were otherwise comparable to individuals in traditional insurance plans, and thus any differences in costs reflect the effects of the HMO on medical utilization. This study found that the cost of medical care for the HMO enrollees was only 72% of the cost for enrollees in traditional insurance plans, in large part because the HMO admitted patients to hospitals at a much lower rate (7.1% of patients were admitted to hospitals by HMOs, compared to 11.2% in the traditional plans).⁵¹ Thus, the cost savings to HMOs do not purely reflect selection (the fact that healthier people enroll in HMOs).

Quality

As we saw earlier, retrospective reimbursement of providers may lead to excessive provision of care. We also saw that prospective reimbursement under HMOs might lead to *underprovision* of care. Do HMOs underprovide or do they simply serve to correct some of the natural excesses of retrospective reimbursement? The literature on the impact of HMOs on patient treatment, and the answer is a definite maybe. Roughly equal numbers of studies find that HMOs deliver care that is better, worse, or the same as traditional health insurance. At this point, there is no real consensus on the impact of HMOs on care quality. While it is clear that HMOs are paying lower prices for medical services, it is not clear that they are actually providing significantly fewer of those services or having any measurable adverse effects on health outcomes.

How Should Providers Be Reimbursed?

The managed care "revolution" does not appear to have had the negative effect on patient health that some feared. The advent of managed care has clearly lowered reimbursement to providers, but it has not measurably lowered the quality of care those providers deliver. Thus, the move from retrospective to prospective reimbursement appears to have improved efficiency in the health care sector. The key question for the future is whether additional "tightening" of the prospective

reimbursement system is needed. We discuss this issue at length when we discuss the Medicare program in the next chapter.

DO NOT COPY
reibaker@live.unc.edu

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DONOTCOPY
reibaker@live.unc.edu

Quality

As we saw earlier, retrospective reimbursement of providers may lead to excessive provision of care. We also saw that prospective reimbursement under HMOs might lead to *underprovision* of care. Do HMOs underprovide or do they simply serve to correct some of the natural excesses of retrospective reimbursement? The literature on the impact of HMOs on patient treatment, and the answer is a definite maybe. Roughly equal numbers of studies find that HMOs deliver care that is better, worse, or the same as traditional health insurance. At this point, there is no real consensus on the impact of HMOs on care quality. While it is clear that HMOs are paying lower prices for medical services, it is not clear that they are actually providing significantly fewer of those services or having any measurable adverse effects on health outcomes.

How Should Providers Be Reimbursed?

The managed care “revolution” does not appear to have had the negative effect on patient health that some feared. The advent of managed care has clearly lowered reimbursement to providers, but it has not measurably lowered the quality of care those providers deliver. Thus, the move from retrospective to prospective reimbursement appears to have improved efficiency in the health care sector. The key question for the future is whether additional “tightening” of the prospective reimbursement system is needed. We discuss this issue at length when we discuss the Medicare program in the next chapter.

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
reibaker@live.unc.edu

Quality

As we saw earlier, retrospective reimbursement of providers may lead to excessive provision of care. We also saw that prospective reimbursement under HMOs might lead to *underprovision* of care. Do HMOs underprovide or do they simply serve to correct some of the natural excesses of retrospective reimbursement? The literature on the impact of HMOs on patient treatment, and the answer is a definite maybe. Roughly equal numbers of studies find that HMOs deliver care that is better, worse, or the same as traditional health insurance. At this point, there is no real consensus on the impact of HMOs on care quality. While it is clear that HMOs are paying lower prices for medical services, it is not clear that they are actually providing significantly fewer of those services or having any measurable adverse effects on health outcomes.

How Should Providers Be Reimbursed?

The managed care “revolution” does not appear to have had the negative effect on patient health that some feared. The advent of managed care has clearly lowered reimbursement to providers, but it has not measurably lowered the quality of care those providers deliver. Thus, the move from retrospective to prospective reimbursement appears to have improved efficiency in the health care sector. The key question for the future is whether additional “tightening” of the prospective reimbursement system is needed. We discuss this issue at length when we discuss the Medicare program in the next chapter.

15.4 Conclusion

This chapter provided an overview of the health care economy in the United States. We began with a discussion of the nature of health insurance in the United States and learned that most individuals have private health insurance and that for those employed by large firms, this is a well-functioning insurance market. For small firms and individuals, there are more failures in the insurance market, which is one possible reason that approximately 26 million Americans are uninsured.

The benefits of health insurance are clear from the theory presented in [Chapter 12](#): risk-averse individuals greatly value the consumption-smoothing benefits of having their medical bills paid. There are clear moral hazard costs as well, both on the patient side and on the provider side. Some cost sharing has been used to address moral hazard on the patient side, and managed care has arisen as a means of addressing moral hazard on the provider side. The success or failure of these approaches is not yet fully apparent. In the next chapter, we discuss how the government has approached the trade-off between providing insurance for those who need it and inducing excess medical care through moral hazard.

HIGHLIGHTS

- Health care has improved dramatically in the United States over the past 70 years, but substantial inequalities remain.
- Most individuals have private insurance, which is largely provided through firms because of risk pooling and the tax subsidy to employees for employer-provided health insurance.
- The 26 million uninsured persons concern policy makers for these reasons: externalities, labor market inefficiencies (job lock), and paternalism.
- There is a clear potential for moral hazard on the patient side of the medical care system because medical demand has been shown to be somewhat elastic, with few health improvements attributable to more generous health insurance plans.
- Despite this, individuals still demand very generous health insurance coverage in the United States, which could be due to the tax subsidy to insurance, access to the most expensive treatments that insurance provides, or psychological motivations.
- The potential for moral hazard on the provider side of the medical care system has given rise to managed care, which uses shopping and prospective reimbursement to control medical costs.
- Existing research suggests that managed care has controlled costs, with no clear evidence that the quality of medical care has suffered as a result.

QUESTIONS AND PROBLEMS

1. Every year the U.S. Census Bureau publishes a report called “Health Insurance Coverage in the United States,” with graphics and statistics on insurance coverage for Americans, using data from the American Community Survey and the Current Population Survey. Navigate to the PDF report for the most recent year available. This can be found in the Census’ “Publications,” at <https://www.census.gov/topics/health/health-insurance/library/publications.html>, or by searching the web for “Health Insurance Coverage in the United States.” As of this writing, the most recent report, published in 2020 and covering the period up to year 2019, was found at

<https://www.census.gov/content/dam/Census/library/publications/2020/demo/p60-271.pdf>.

- a. Find a graphic depicting the U.S. uninsured rate over time for different age groups. In what year did the largest change in the uninsured rate occur? For which age groups was the effect largest? What explains this change? Prior to that year, what was the trend in the uninsured rate? What was the trend after that year?
 - b. What states have the highest and lowest uninsured rates? What states have seen the greatest drop in the uninsured rate over the past several years? What states have seen the greatest increase?
2. Suppose the U.S. government eliminates the tax exemption for employer-provided health insurance. Instead, the government provides a 20% subsidy on employer-provided health insurance, so that the employer only has to pay 80% of the cost of such policies. How might this new policy affect the type of workers to whom firms will offer health insurance? What types of firm is this policy most likely to affect?
 3. Prior to the Affordable Care Act, many privately purchased non-group health insurance plans had stringent “preexisting condition” exclusions, which denied coverage to insured persons for any health conditions that already existed at the time of enrollment.
 - a. How did these preexisting condition exclusions arise from a market failure in the insurance market?
 - b. How did the ACA change the treatment of preexisting conditions in the private-purchase non-group health insurance market?
 4. What negative externalities arise when an individual does not have health insurance?
 5. An individual's demand for physician office visits per year is
$$Q = 10 - (1/20)P$$
, where P is the price of an office visit. The marginal cost of producing an office visit is \$120.
 - a. If individuals pay full price for obtaining medical services, how many office visits will they make per year?
 - b. If individuals must pay only a \$20 copayment for each office visit, how many office visits will they make per year?
 - c. What is the deadweight loss to society associated with not charging individuals for the full cost of their health care?
 6. Aditya has three types of medical expenditures: prescription eyeglasses, adalimumab, a brand-name anti-rheumatic drug that he takes twice monthly for rheumatoid arthritis, and accidents and future illnesses (such as broken bones, pneumonia, or cancer). He has been paying for all of his medical expenditures out of pocket and he is now considering purchasing health insurance. Different plans he is considering offer coverage for

Health insurance. Different plans we're considering offer coverage for

different types of expenditures, including vision, prescription drug, and hospital benefits. Describe the consumption-smoothing benefits and moral hazard costs of coverage for prescription eyeglasses, anti-rheumatic drugs, and accidents and future illnesses.

7. [Figure 15-1](#) shows that the United States spends the most on health care of all OECD countries, with health expenditures as a percentage of GDP approximately double the OECD average. Despite this higher level of expenditure, [Figure 15-3](#) shows that other countries outperform the United States on most health care outcomes. Discuss several factors that might contribute to this disconnect between spending and outcomes.
8. **e** Senator Lee, making the case for universal, free health care, argues that we don't need to worry that health care costs might go up under such a universal and free system, because people are not price sensitive when it comes to their health. She argues that when people need to go to the doctor, they go, regardless of the cost. Evaluate her argument in light of the empirical evidence on the price sensitivity of health care demand. According to research discussed in the chapter, to what degree do you think health care spending might change under a universal, free health care system?
- 9.

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DONOTCOPY
reibaker@live.unc.edu

CHAPTER 16

Health Insurance II: Medicare, Medicaid, and Health Care Reform



James Andrews/Stock/Gett Images

[16.1 The Medicaid Program for Low-Income Families](#)

[16.2 What Are the Benefits of the Medicaid Program?](#)

[16.3 The Medicare Program](#)

[16.4 Controlling Costs in the Medicare Program](#)

[16.5 Long-Term Care](#)

[16.6 Health Care Reform and the ACA](#)

[16.7 Conclusion](#)

Questions to keep in mind

- How does the Medicaid program impact patient health?
 - How have cost control efforts fared in the Medicare program?
 - How does the Affordable Care Act work?
-

Fundamental reform of the U.S. health care system has been a constant source of political debate for much of the past century. In the postwar period, there were major attempts to transform health care by President Harry S. Truman in 1950, President Richard Nixon in 1974, and President Bill Clinton in 1994. All failed, in some cases dramatically. The losing streak for health care reform ended in 2010,

however, when President Barack Obama signed into law the most sweeping overhaul of the U.S. health care system since the introduction of Medicare and Medicaid in 1965.

The Patient Protection and Affordable Care Act (ACA) made four fundamental changes to the U.S. health care system.

- It banned insurers from denying coverage because of pre-existing conditions and from charging different prices to different people based on their health.
- It mandated that all residents of the United States be covered by health insurance, with exemptions for illegal immigrants, very-low-income families, and those for whom insurance is unaffordable.
- It required the federal government to extensively subsidize health insurance coverage for the poor by expanding the Medicaid program and by implementing sizeable new tax credits that offset insurance costs.
- It took a variety of actions to lower health care costs, including the introduction of private insurance exchanges through which insurers face competitive pressure to lower their prices, a cap on the tax exclusion for employer-provided insurance, and alternative provider reimbursement arrangements designed to move toward the optimal provider reimbursement strategies discussed in [Chapter 15](#).

This legislation was highly controversial and passed through Congress with a very slim margin in a strictly partisan vote (no Republicans voted for it). Critics on the right worried that the law would lead to restricted patient choice and a bloated government bureaucracy. As Republican presidential candidate Herman Cain said in August 2011, “If ObamaCare had been fully implemented when I caught cancer, I’d be dead.”¹

But those on the left were not particularly pleased either: they believed that this proposal represented a retreat from the government-run [single-payer system](#) that might more efficiently expand coverage and control costs. As former Vermont governor Howard Dean said, “This is essentially the collapse of health care reform in the U.S. Senate.” This debate between the right and the left highlights the critical argument about the proper role of the government in the U.S. health care system.

single-payer system

A health care system in which all health insurance is provided and paid for by the government.

Perhaps due to the dissatisfaction from both sides of the aisle, the Affordable Care Act struggled to gain popularity in the United States, and opposition to the ACA was a rallying cry for Republican candidates in the 2016 elections. President Trump entered office with a Republican congressional majority in January of 2017, and the future of the ACA looked dire.

But then the tables turned. The ACA's popularity began to grow as more Americans benefited from its provisions. By 2017, an estimated twenty million Americans had gained coverage through the ACA, mostly because of its expansion of public insurance through the Medicaid program. Meanwhile, health care costs continued to grow, but more slowly than in the past, and economic growth continued unabated, weakening concerns that the ACA would lead to runaway health care inflation and reduced economic activity. As a result, when the Republican Congress introduced repeated bills to repeal the ACA during the 2017-2018 session, they were unable to pass. A series of subsequent legislative and regulatory actions, most notably the repeal of the individual mandate, have significantly weakened the law, but it remains the law of the land.

Despite the general perception of the United States as having a private health care system, the government plays an enormous role in the provision of health care; almost half of all health care spending today is done by the government.² Public health spending is targeted to three specific groups: the Medicare program targets older adults and people with disabilities, and the Medicaid program targets low-income families, including low-income elderly and disabled individuals who do not qualify for Medicare. Medicaid and Medicare represent the most rapidly growing expenditure item for federal and state governments; the programs did not exist until 1965 but today account for almost a quarter of government spending.³ Because of their large and growing role, these programs are a constant source of policy debate.

To help you become familiar with the functions of these programs, this chapter provides a description of how and for whom they operate. We also review the major policy issues concerning these two programs. We begin by assessing Medicaid's impact on the insurance coverage and health of low-income families participating in the program: Has expanding public insurance to low-income groups proved a cost-effective means of improving their health? We then turn to a major source of policy debate: how to control the rapidly rising costs of the Medicare program. We discuss the two major government efforts to control costs: moving to prospective reimbursement of providers and enabling consumer choice of managed care plans.

Next, we discuss the set of issues surrounding long-term care of older adults, which involves both Medicaid and Medicare programs.

Finally, the last section of the chapter pulls together the lessons from [Chapter 15](#) and this chapter to discuss the broad question of health care reform in the United States and to discuss the design of the ACA and early evidence on its effects.

DONOTCOPY
reibaker@live.unc.edu

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
reibaker@live.unc.edu

To help you become familiar with the functions of these programs, this chapter provides a description of how and for whom they operate. We also review the major policy issues concerning these two programs. We begin by assessing Medicaid's impact on the insurance coverage and health of low-income families participating in the program: Has expanding public insurance to low-income groups proved a cost-effective means of improving their health? We then turn to a major source of policy debate: how to control the rapidly rising costs of the Medicare program. We discuss the two major government efforts to control costs: moving to prospective reimbursement of providers and enabling consumer choice of managed care plans. Next, we discuss the set of issues surrounding long-term care of older adults, which involves both Medicaid and Medicare programs.

Finally, the last section of the chapter pulls together the lessons from [Chapter 15](#) and this chapter to discuss the broad question of health care reform in the United States and to discuss the design of the ACA and early evidence on its effects.

16.1 The Medicaid Program for Low-Income Families

The major public health intervention for low-income populations in the United States is the Medicaid program. This program serves two types of groups: low-income families and low-income disabled and elderly individuals. Because the majority of program enrollees are in the first group, our discussion focuses on low-income families; there will be a brief discussion of people with disabilities and older adults later in the chapter.

How Medicaid Works

Medicaid, like unemployment insurance (UI), is a program that is federally mandated but administered by states.⁴ The program is financed by matching grants from the federal government, paid out of general revenues. The rate at which the federal government shares in Medicaid spending is an inverse function of state income, with the federal government paying half the costs of the program in high-income states, such as Massachusetts, and more than three-quarters in low-income states, such as Mississippi. The federal government mandates some minimal levels of eligibility and service coverage, and states are free to increase generosity beyond these mandates.

Individuals qualify for Medicaid on the basis of their income and family structure, as described next. If eligible, individuals may then enroll in the Medicaid program. Medicaid insurance coverage is similar to private insurance: it reimburses providers for the services they provide to enrollees, although there is little or no patient coinsurance.

Who Is Eligible for Medicaid?

Medicaid was introduced in the late 1960s as a health insurance component of state cash welfare programs that targeted low-income, single-parent families. Beginning in the mid-1980s, the Medicaid program was slowly separated from cash welfare programs, first by extending benefits to low-income children in two-parent families and then by raising the income eligibility thresholds for two groups—children and pregnant women (who were covered only for the costs associated with pregnancy, not for other health costs).

In 1997, the Medicaid program for children was augmented by the **Children's Health Insurance Program (CHIP)**. The goal of CHIP was to expand the eligibility of children for public health insurance beyond the existing limits of the Medicaid program, using either expansions of the Medicaid program or a new program that more closely mimics private health insurance. To provide incentives for states to expand their low-income health care coverage using CHIP funds, the federal government pays a higher share of each state's CHIP costs than it pays of the state's Medicaid costs. In the remainder of this chapter, we will typically refer to both Medicaid and CHIP coverage as *Medicaid*.

Children's Health Insurance Program (CHIP)

A program introduced in 1997 to expand eligibility of children for public health insurance beyond the existing limits of the Medicaid program, generally up to 200% of the poverty line.

Currently, all individuals age 18 or younger are eligible for Medicaid or CHIP up to 100% of the poverty line (\$26,200 for a family of four), and children under age 6 and pregnant people are covered up to 138% of the poverty line (\$36,160). In most states, eligibility extends further for both children and pregnant people a typical state covers both groups up to 200% of the poverty line (\$52,400). Eligibility extends even further in some states: pregnant people in Iowa, for example, are eligible up to 375% of the poverty line (\$98,250).⁵

More recently, there was a massive expansion of the Medicaid program put in place by the ACA, which extended eligibility to all families, regardless of family structure, below 138% of the federal poverty line.⁶ This expansion was designed to cover the entire United States, but due to a Supreme Court decision in 2012, states were able to opt out of the expansion. Even though the federal government is paying at least 90% of the costs of the expansion, as of 2020, 12 states have opted out and therefore not expanded Medicaid to this new group.⁷

What Health Services Does Medicaid Cover?

Besides setting the rules for Medicaid eligibility, states also control two other aspects of the Medicaid program. The first is service coverage. While federal Medicaid rules require states to cover major services, such as physician and hospital care, they do not require states to pay for optional services, such as prescription drugs or dental care. Even so, all states have chosen to cover the most expensive optional benefits; all states cover prescription drugs, for example, and all but two cover at least a portion of dental services.⁸ For the traditional Medicaid population, these services are provided with little or no copayment required. (In states that

have CHIP, copayments are allowed to be somewhat higher for those above 150% of the poverty line.) This package of services is much more generous than that of virtually any private insurance plan. Thus, Medicaid may be the best insurance money can't buy.

How Do Providers Get Paid?

States can also regulate the rate at which health service providers are reimbursed. Unlike the situation for services covered (in which all states cover basically the same health care services), there is more variability across the states in provider reimbursements. In most states, Medicaid reimburses physicians at a much lower level than does the private sector, which often leads physicians to be unwilling to serve Medicaid patients. For childbirth, for example, the reimbursement rate to physicians under Medicaid averages about half of the private-sector reimbursement rate for the same services. Perhaps as a result, nearly one-third of physicians reported in a 2011 survey that they would not accept new Medicaid patients.⁹ Thus, while the coverage provided by Medicaid is very generous in all states, individuals may have trouble availing themselves of that coverage in a number of states because physicians do not want to accept them as patients.¹⁰

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DONOTCOPY
reibaker@live.unc.edu

16.2 What Are the Benefits of the Medicaid Program?

The Medicaid program is not only enormous, with spending of \$613.5 billion in 2019, but it is growing rapidly—Medicaid spending is expected to increase at a rate of 5.7% on average over the next several years.¹¹ The goal of this large and rapidly growing program is to provide health insurance coverage to low-income populations who cannot afford private coverage, thus providing them with financial protection and ultimately improving their health. Whether Medicaid has had the intended effects is an empirical question, and in this section, we review the evidence on this question.

Does Medicaid Provide Financial Protection?

The primary goal of insurance programs such as Medicaid is to provide financial protection against the risks imposed by medical spending. This is particularly important for the very-low-income families who qualify for this program, for whom even a modest health care bill can exceed their financial resources.

There are a number of studies that have documented the improved financial protection provided by Medicaid. [Gruber and Yelowitz \(1999\)](#) show that expansions of the Medicaid program are associated with higher levels of consumption on nonmedical goods. [Gross and Notowidigdo \(2011\)](#) show that expanding the Medicaid program lowers the rate of consumer bankruptcy. And [Finkelstein et al. \(2012\)](#) show that individuals made eligible for Medicaid in the state of Oregon see dramatic reductions in their levels of out-of-pocket medical spending, medical debt, and bills sent to collection agencies. Clearly, Medicaid *is* providing important consumption smoothing for low-income families.

Does Medicaid Improve Health?

While the impact of public insurance on financial protection is fairly straightforward, the impact on health is more difficult to assess. [Figure 16-1](#) provides an organizing framework for answering this question, in several steps.

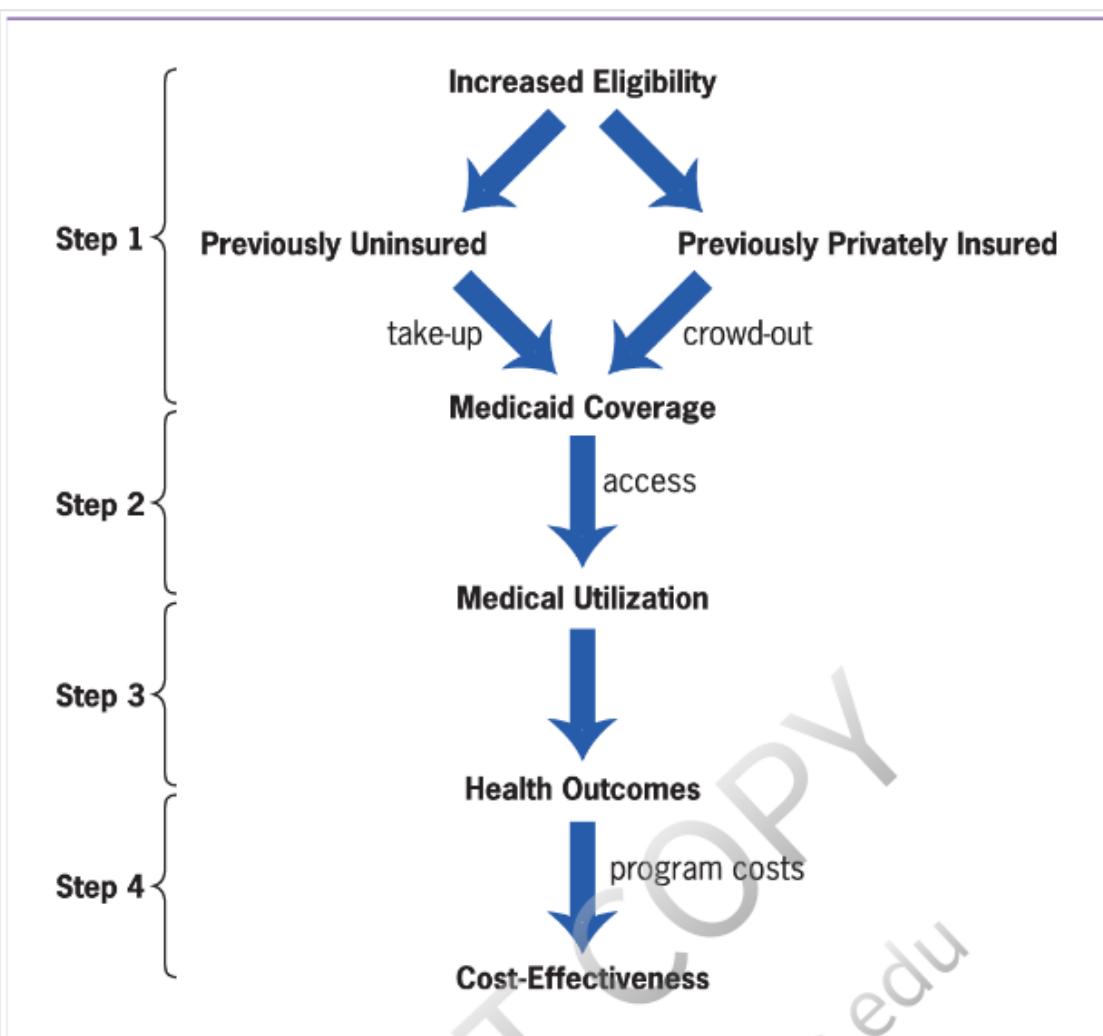


FIGURE 16-1 How Does Medicaid Affect Health? • This diagram shows the channels through which expanding eligibility for Medicaid can affect health. Expansions increase eligibility and lead to greater take-up by the uninsured and some crowd-out of private insurance (Step 1). This increase in insurance coverage leads to increased medical utilization (Step 2), depending on access restrictions for Medicaid enrollees. Higher medical utilization then potentially leads to better health (Step 3). This better health is associated with increased program costs, leading to the ultimate question of the cost-effectiveness of this route to improving health (Step 4).



Step 1: Translate the eligibility change into a change in actual coverage by the Medicaid program. There are two channels through which eligibility increases can increase coverage. The first is the “take-up” channel: eligibility changes make previously uninsured people eligible for coverage, but only some of these newly eligible people will enroll in the program. The second is the “crowd-out” channel: eligibility causes some people with private insurance to discontinue their private insurance and enroll in public insurance. This outcome is another form of the substitution of social insurance for self-insurance (crowd-out) that we discussed in [Chapters 12, 13, and 14](#).

Step 2: Translate the increases in Medicaid coverage into actual increases in utilization of care. As noted earlier, many providers will not see Medicaid patients because of low reimbursement rates. As a result, it is not clear that covering more people with Medicaid will greatly increase utilization; if the supply of providers willing to see Medicaid enrollees is fixed, increasing demand for care will not increase the total amount of care provided.

Step 3: Translate the increase in utilization of care into improved health. It is not clear how important increased medical care will be for health improvements. The RAND Health Insurance Experiment (HIE) results reviewed in the previous chapter, for example, suggest that varying the generosity of health insurance does not improve health outcomes very much.

Step 4: Determine the cost-effectiveness of the Medicaid expansion. The value of any Medicaid-induced improvements in health must be weighed against their costs. Even if Medicaid improves health, if we could improve health more cost-effectively through other policies, then we may want to pursue those other policies instead.

How Does Medicaid Affect Health? Evidence

What is the evidence concerning the various steps shown in [Figure 16-1](#)? In this section, we provide a quick overview of the relevant economics literature.

Take-Up

The Medicaid expansions of the 1980s and 1990s led to an explosion in eligibility for the program. In 1982, 12% of individuals nationwide aged 18 or under were eligible for public insurance under Medicaid. By 2000, 46% of individuals in that age group were eligible, an increase of almost 400%. There was a parallel rise for pregnant people with some small increases for parents of eligible children in selected states that chose to expand to that population.^{[12](#)}

Relatively few of the newly eligible people enrolled, however. By most estimates, only about one-quarter of those made eligible for Medicaid in the late 1980s and early 1990s enrolled in the program, and only about 10% of those made eligible through the CHIP expansions of the late 1990s enrolled. Just as we saw with UI in [Chapter 14](#), some of this low enrollment may be caused by a lack of information about eligibility for the program, and some may be caused by stigma about taking a public handout. Another reason that is particular to the Medicaid expansions, however, is that most of those made eligible *already had private health insurance coverage*. For those individuals, switching from private to public coverage might not

be very attractive; for example, individuals may be wary of leaving their private insurance for a public program from which they could be disqualified if their incomes rise.

Crowd-Out

Unlike people who prefer to hold on to their private health insurance, some individuals might find it attractive to leave private insurance for public insurance because the Medicaid insurance package is much more generous than that provided by the typical private insurance plan and it doesn't cost anything. In terms of benefits, many employer-provided insurance plans don't include the "optional" benefits covered by most state Medicaid programs, such as dental or vision care. In terms of costs, the typical employer that offers health insurance charges families more than \$477 per month, or \$5,726 per year, to enroll in an insurance plan, while Medicaid is free.¹³ Moreover, while copayments are low in employer-provided insurance relative to the optimal level recommended by economic theory, they are much higher than those in Medicaid, which has close to zero copayments as mandated by law. As a result of all these cost and coverage differences between private and public insurance, some privately insured individuals may find it advantageous to switch to public insurance when they become eligible. This is another example of the ways in which government intervention can crowd out private provision, as we have seen with fireworks, education, and social insurance.

There are a number of empirical studies that estimate the extent of crowd-out. These studies have found that there is some crowd-out, but it is far from complete, with private insurance declines amounting to 20 to 50% of the public insurance increases.¹⁴

Health Care Utilization and Health

Even at the largest estimates of crowd-out, expanding Medicaid still substantially reduces the number of uninsured, so expansions may affect the utilization of health care services (such as doctor's office visits) and ultimately the health of those newly covered. In fact, the existing evidence suggests that both utilization of health care services and health care outcomes improved when Medicaid expanded (Steps 2–3 in [Figure 16-1](#)). Preventive care, particularly early prenatal care and preventive medical visits by children, rose by more than 50% when individuals were made eligible for Medicaid. At the same time, there were large corresponding reductions in both infant and child mortality: infant mortality, for example, declined by 8.5% as a result of the expansion of Medicaid coverage to pregnant people.¹⁵

The finding that providing health insurance to the uninsured can improve their health is echoed in other studies as well. [Hanratty \(1996\)](#) studied the introduction of national health insurance in Canada and found that it was associated with a 4% decline in the infant mortality rate and an 8.9% decrease in the incidence of low birth weight among infants of single parents. [Lurie et al. \(1984\)](#) studied a large group of individuals in California who lost their eligibility to receive public insurance due to a fiscal crisis in the early 1980s that forced the state to cut back its insurance coverage and found that health deteriorated significantly among this group afterward, including a higher rate of death. A study of free universal primary care provision in Turkey found that it reduced mortality rates by more than 25% among infants, 23% among those age 1–4, and 7% among older adults.¹⁶ A recent study of removing access to public health care for undocumented immigrants in Spain found that it led to a significant increase in mortality among that group.¹⁷ Finally, a study by [Finkelstein et al. \(2012\)](#) looked at a novel randomized trial of health insurance coverage that operated similarly to the school choice lotteries discussed in [Chapter 11](#). The state of Oregon had funding for 10,000 new spots in its public insurance program, but when it offered individuals the chance to enroll, almost 100,000 signed up. The state therefore allocated the spots randomly by lottery. This study found that lottery winners were not in much better health physically, but saw large improvements in mental health: there was a huge reduction in reported depression, at least partly due to the reduced stress of self-financing one's health care.

Do these findings contradict the conclusions of the RAND HIE, which found that varying the extent of insurance (through the coinsurance rate) had no impact on health? They don't, because the HIE made no one uninsured; all individuals were fully insured for expenditures above \$1,000. So putting these two pieces of evidence together, we can trace the type of medical effectiveness curve that we showed in [Figure 15-8](#). Moving individuals from being uninsured to having some insurance, as is achieved by the Medicaid expansions, has an important positive effect on health. But once someone is insured, varying the amount of insurance coverage, as is done in the RAND HIE, doesn't seem to cause significant changes in health.

Cost-Effectiveness

Finally, we come to cost-effectiveness, the last arrow in [Figure 16-1](#); evidence indicates that expanding public insurance does improve health, but at what cost? [Currie and Gruber \(1996a\)](#) estimated that it cost Medicaid roughly \$1 million per infant life saved through its expansions. This is much lower than the typical

estimate of the value of a life from compensating differential studies discussed in [Chapter 8](#) (\$3 million to \$7 million). That \$1 million is also much lower than the cost of many alternative government interventions designed to save lives, such as food regulation or seat-belt safety. This finding suggests that investing in low-income health care may be a cost-effective means of improving health in the United States.

EMPIRICAL EVIDENCE

Using State Medicaid Expansions to Estimate Program Effects

A natural way to measure the effect of Medicaid on health status would be to compare those who choose to enroll in the program with those who are not enrolled. Such an approach suffers from two types of bias, however. First, eligibility for Medicaid is determined by factors, such as income, that might also determine outcomes such as health status (e.g., those with incomes low enough to be eligible for Medicaid may also be the least healthy individuals in a state). Second, only some individuals who are eligible for Medicaid will take up the program, and these individuals may be different from those who do not take up the program (e.g., those enrolled may be in worse health). As a result, the treatment group (those on Medicaid) will be different along many dimensions from the control group (those not on Medicaid), biasing estimates of Medicaid's effect on health outcomes.

Fortunately, the expansions in Medicaid eligibility over the 1980s and 1990s provide a natural means for addressing these biases. An important feature of the Medicaid expansions is that they occurred at a very different pace across the states and at a different pace for different age groups of children within states, as illustrated in [Table 16-1](#). In Missouri, for example, which had fairly restrictive eligibility for Medicaid before these expansions, Medicaid eligibility for children rose from 12% of children in 1982 to 76% of children in 2000. In Michigan, in contrast, 20% of children were already eligible in 1982, and that figure rose to only 34% by 2000.

TABLE 16-1 Medicaid Eligibility Changes Across and within States

Eligibility for all children, by state		
Year	Missouri eligibility	Michigan eligibility
1982	12%	20%
2000	76%	34%
Eligibility for children by age, in Washington, D.C.		
Year	Age 13 eligibility	Age 0 eligibility
1982	18%	48%
2000	59%	56%

Data from: Calculations from the authors research with Kosali Simon at Cornell University.

Over the 1982–2000 period, Medicaid eligibility rose much more in Missouri than in Michigan (top panel). There were also dramatic differences in eligibility growth within states: eligibility rose much more for 13-year-olds in Washington, D.C., than for 0-year-olds (bottom panel).

This differential pace of expansion across the states provides an excellent setting for quasi-experimental analysis of the effects of these programs. Studies can compare outcomes (such as degree of illness) in the treatment states (those that expanded eligibility more) to outcomes in the controls (those that expanded it less). As long as nothing else is changing in these states that is correlated with both the eligibility expansions and the outcome variables, such as health, then this approach controls for the bias inherent in comparing individuals on and off Medicaid.

It is possible, however, that other things were changing along with state insurance expansions. For example, states may have been more willing to expand Medicaid when there was a state recession, which would independently affect health (if parents have less money, they are less able to afford food and health services for their children). As in the case of Unemployment Insurance, however, there are "within-state" control groups that can further limit bias because eligibility expands for some age groups of children and not for others. For example, researchers could compare what happened to outcomes of 13-year-olds in Washington, D.C., the treatment group, for whom eligibility rose by more than 40% of the population from 1982 to 2000 (from 18 to 59%), with what happened to newborns there, the control group, for whom eligibility rose by less than 10% of the population (from 48 to 56% because a large percentage of newborns were already eligible in 1982). Both groups were subject to the same outside factors, such as a recession, but they faced very different Medicaid policies. As shown in the appendix to [Chapter 14](#), we can combine the across-state comparisons with the within-state comparisons to develop a more convincing quasi-experimental estimate.

The more recent Medicaid expansions under the ACA provide a new quasi-experiment for measuring the effects of health insurance, comparing states that did and did not expand Medicaid to individuals below 133% of the poverty line. A large literature has emerged that uses these recent expansions to confirm the findings from the earlier literature: expanding Medicaid leads to large increases in access to care, utilization of care, and improvements in health.^{[18](#)}

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
reibaker@live.unc.edu

Cost-Effectiveness

Finally, we come to cost-effectiveness, the last arrow in [Figure 16-1](#); evidence indicates that expanding public insurance does improve health, but at what cost? [Currie and Gruber \(1996a\)](#) estimated that it cost Medicaid roughly \$1 million per infant life saved through its expansions. This is much lower than the typical estimate of the value of a life from compensating differential studies discussed in [Chapter 8](#) (\$3 million to \$7 million). That \$1 million is also much lower than the cost of many alternative government interventions designed to save lives, such as food regulation or seat-belt safety. This finding suggests that investing in low-income health care may be a cost-effective means of improving health in the United States.

EMPIRICAL EVIDENCE

Using State Medicaid Expansions to Estimate Program Effects

A natural way to measure the effect of Medicaid on health status would be to compare those who choose to enroll in the program with those who are not enrolled. Such an approach suffers from two types of bias, however. First, eligibility for Medicaid is determined by factors, such as income, that might also determine outcomes such as health status (e.g., those with incomes low enough to be eligible for Medicaid may also be the least healthy individuals in a state). Second, only some individuals who are eligible for Medicaid will take up the program, and these individuals may be different from those who do not take up the program (e.g., those enrolled may be in worse health). As a result, the treatment group (those on Medicaid) will be different along many dimensions from the control group (those not on Medicaid), biasing estimates of Medicaid's effect on health outcomes.

Fortunately, the expansions in Medicaid eligibility over the 1980s and 1990s provide a natural means for addressing these biases. An important feature of the Medicaid expansions is that they occurred at a very different pace across the states and at a different pace for different age groups of children within states, as illustrated in [Table 16-1](#). In Missouri, for example, which had fairly restrictive eligibility for Medicaid before these expansions, Medicaid eligibility for children rose from 12% of children in 1982 to 76% of children in 2000. In Michigan, in contrast, 20% of children were already eligible in 1982, and that figure rose to only 34% by 2000.

TABLE 16-1 Medicaid Eligibility Changes Across and within States

Eligibility for all children, by state		
Year	Missouri eligibility	Michigan eligibility
1982	12%	20%
2000	76%	34%
Eligibility for children by age, in Washington, D.C.		
Year	Age 13 eligibility	Age 0 eligibility
1982	18%	48%

2000	59%	56%
Data from: Calculations from the authors research with Kosali Simon at Cornell University.		
Over the 1982–2000 period, Medicaid eligibility rose much more in Missouri than in Michigan (top panel). There were also dramatic differences in eligibility growth within states: eligibility rose much more for 13-year-olds in Washington, D.C., than for 0-year-olds (bottom panel).		

This differential pace of expansion across the states provides an excellent setting for quasi-experimental analysis of the effects of these programs. Studies can compare outcomes (such as degree of illness) in the treatment states (those that expanded eligibility more) to outcomes in the controls (those that expanded it less). As long as nothing else is changing in these states that is correlated with both the eligibility expansions and the outcome variables, such as health, then this approach controls for the bias inherent in comparing individuals on and off Medicaid.

It is possible, however, that other things were changing along with state insurance expansions. For example, states may have been more willing to expand Medicaid when there was a state recession, which would independently affect health (if parents have less money, they are less able to afford food and health services for their children). As in the case of Unemployment Insurance, however, there are “within-state” control groups that can further limit bias because eligibility expands for some age groups of children and not for others. For example, researchers could compare what happened to outcomes of 13-year-olds in Washington, D.C., the treatment group, for whom eligibility rose by more than 40% of the population from 1982 to 2000 (from 18 to 59%), with what happened to newborns there, the control group, for whom eligibility rose by less than 10% of the population (from 48 to 56% because a large percentage of newborns were already eligible in 1982). Both groups were subject to the same outside factors, such as a recession, but they faced very different Medicaid policies. As shown in the appendix to [Chapter 14](#), we can combine the across-state comparisons with the within-state comparisons to develop a more convincing quasi-experimental estimate.

The more recent Medicaid expansions under the ACA provide a new quasi-experiment for measuring the effects of health insurance, comparing states that did and did not expand Medicaid to individuals below 133% of the poverty line. A large literature has emerged that uses these recent expansions to confirm the findings from the earlier literature: expanding Medicaid leads to large increases in access to care, utilization of care, and improvements in health.¹⁸

16.3 The Medicare Program

The largest public health insurance program in the United States is Medicare, which, as previously discussed, was started in 1965 as a universal health insurance system for the older adults and people with disabilities under age 65 (those receiving disability insurance from the federal government). To help you organize your thoughts about Medicaid and Medicare, [Table 16-2](#) compares the key features of the two programs.

TABLE 16-2 Medicaid and Medicare as of 2021

	Medicaid	Medicare
Eligibles	Families on welfare Low-income children, pregnant people Low-income older adults, people with disabilities	Retirees and spouses 65 and older Certain people with disabilities under 65 People with kidney failure (requiring dialysis or transplant)
Premiums	None	Hospital coverage: none Physician coverage: \$148.5 per month Prescription drug coverage: Variable
Deductibles/copayments	None (or very small)	Hospital coverage: \$1,484 deductible for benefit period Physician coverage: \$203 deductible, 20% coinsurance Prescription drug coverage: Variable
Services excluded	None (or very minor)	Prescription drugs (until 2006), routine checkups (until 2010), dental care, nursing home care, eyeglasses, hearing aids
Provider reimbursement	Very low	Moderate (but falling)

Medicaid provides health insurance for low-income individuals, covering a wide range of health services at little cost to those individuals. Medicare provides health insurance for those age 65 and over, covering many, though not all, health services at some cost to those individuals.

How Medicare Works

The Medicare program is administered at the federal level. All U.S. citizens who have worked and paid payroll taxes for ten years, and their spouses, are eligible for coverage; other citizens who do not have the requisite work experience can purchase Medicare coverage at its full cost. Medicare operates similarly to a private insurance plan, with the government reimbursing providers for their costs and with patients responsible for coinsurance. There are two key features of Medicare to keep in mind as we learn about this program and think about reforms to it.

Medicare Is Really Three Different Programs

Medicare consists of three different programs. **Medicare Part A** covers inpatient hospital costs and some costs of *long-term care* (care for older adults, either in institutions such as nursing homes or in their own homes). This part of the program is financed from the Medicare Hospital Insurance trust fund, which is funded by the payroll tax on both employers and employees discussed in [Chapter 15](#). **Medicare Part B** covers physician expenditures, outpatient hospital expenditures, and other services. About 25% of the cost of this part of the program is financed by enrollee premiums, which are deducted directly from Social Security payments; the remaining 75% of the cost is paid from the general government revenues. Premiums for Part B depend on income; for most elders, they are \$148.50 per month, but they rise to \$475.20 per month for elders with combined incomes above \$330,000.¹⁹ **Medicare Part D** provides coverage for prescription drug expenditures. As described later, individuals can choose from a large variety of private insurance plans for their prescription drug coverage. Part D coverage is financed by a mix of enrollee premiums (which vary widely across available plans) and general revenues.

Medicare Part A

Part of the Medicare program that covers inpatient hospital costs and some costs of long-term care; financed from a payroll tax.

Medicare Part B

Part of the Medicare program that covers physician expenditures, outpatient hospital expenditures, and other services; financed from enrollee premiums and general revenues.

Medicare Part D

Part of the Medicare program that covers prescription drug expenditures.

Medicare Has High Patient Costs

Relative to private health insurance, and certainly relative to Medicaid, the Medicare program has fairly high copayments and deductibles and a relatively lean benefits package. Part A of the program has a \$1,484 deductible for the first 60 days of a hospital stay (with costs to the patient rising to \$371 per day for days 61–90, \$742 per day for days 91–150, and full payment required after that), and Part B has a \$203 deductible and a 20% coinsurance rate.²⁰ It is important to note that the Part B coinsurance is not capped at some level of out-of-pocket expenditures; if an individual has \$10,000 in physician bills in a year, they must pay \$2,000 in coinsurance. This greatly lowers the consumption-smoothing value of Medicare because there is still some risk of very high medical expenditures if you are ill. Part D coverage also features high patient costs (see the Application that follows),

although there is wide variation across Part D insurance plans, with some plans charging higher premiums and in return covering a much larger share of prescription expenditures. Moreover, Medicare does not cover many benefits provided by private-sector insurance plans, including dental and vision care.

APPLICATION

The Medicare Prescription Drug Debate



One of the liveliest health policy debates of the early twenty-first century was over the addition of a prescription drug benefit to the Medicare program. When Medicare was established in 1965, it covered most medical needs for older adults and people with disabilities, including hospital and doctor costs, but it excluded coverage for prescription drugs. This omission was not perceived as a major one in the early years of the Medicare program, but in the 1990s, the advancement of prescription drug treatments for common illnesses among older adults drew attention to this gap in Medicare coverage.

Democrats and Republicans proposed two different approaches to deal with this problem. Democrats suggested adding a drug benefit to the Medicare program, with the government negotiating directly with drug companies to ensure the lowest drug prices. Republicans suggested that the government subsidize private insurers to offer prescription drug coverage to older adults, either through HMOs or as a stand-alone prescription drug-only plan.

There are advantages and disadvantages to both these approaches. On the one hand, the federal government would represent an enormous buying pool for prescription drugs, which would allow it to both minimize administrative costs and strongly negotiate for low prices. On the other hand, the federal government could become too heavy-handed and lower prices too much, thus making new drug development unprofitable and reducing pharmaceutical innovation.²¹ Private insurance approaches would not suffer from this problem because they would negotiate prices with manufacturers, but the private model would introduce the new problem of adverse selection, whereby plans might suffer from getting only those enrollees who need the most prescription drugs.

In December 2003, President George W. Bush signed into law a bill that followed the Republican approach but provided government reinsurance to deal with potential problems of adverse selection. Under this reinsurance arrangement, the federal government reimburses insurers for a share of very large drug bills to ensure that insurers do not suffer unduly from enrolling sick individuals.

Enrollment in the new program, called Part D, was initially fraught with problems. Starting shortly before the official opening of the enrollment period on November 15, 2005, Medicare and state officials were bombarded by hundreds of calls per day from frustrated Medicare recipients who felt hopelessly lost among the plans they were expected to choose from, some of which continued to be adjusted by fiercely competing private insurers even as consumers were trying to make a decision. Despite this rocky start, in the following months, the federal government was able to iron out many of the problems that had arisen during the initial transition, and by the end of 2006, more than 23 million elders were enrolled in the program.²²

How successful has the private-based approach been? The evidence is unclear. On the one hand, the costs of this program have been much lower than projected in the early years. The average monthly

costs of this program have been much lower than projected in the early years. The average monthly premium that enrollees paid under Part D for 2020 was \$32.74, which was 26% lower than the Congressional Budget Office (CBO) projection when the law was passed in 2003.²³ Part of the reason for these lower costs was aggressive competition between insurers and the ability of insurers to steer Part D recipients to lower-cost “generic” drugs ([Duggan and Scott Morton, 2010](#)). On the other hand, there is clear evidence that elders have not been making the choices that are best for them among available insurance options. [Abaluck and Gruber \(2011\)](#) studied a large sample of elders using prescription drugs and found that the typical elder in their sample could have saved 30% by choosing a less expensive plan.

What is clear, however, is that the basic structure of this drug benefit was illogical. For basic Part D plans in 2006, individuals received coverage for:

- None of the first \$250 in drug costs each year
- 75% of costs for the next \$2,250 of drug spending (up to \$2,500 total)
- 0% of costs for the next \$3,600 of drug spending (up to \$5,100 total)
- 95% of costs above \$5,100 of drug spending²⁴

This is a very odd structure, featuring generous coverage for low-spending amounts (between \$250 and \$2,500 of drug spending), followed by a “donut hole” where there is no coverage (up to \$5,100), and then almost full coverage above a catastrophic level (above \$5,100). There is no coherent economic rationale for such a structure. The optimal insurance arrangement discussed in [Chapter 15](#) would feature little coverage of initial spending to avoid moral hazard, with insurance coverage rising as spending rises to ensure consumption smoothing. The first and last brackets of this plan, with an initial deductible and catastrophic coverage above \$5,100, follow that arrangement. Yet there is no economic rationale for the two middle brackets, which provide fairly generous coverage of drug spending, between \$250 and \$2,500, and then no coverage at all for the next \$3,600 of spending. This is an upside-down benefit structure that is exactly the reverse of an optimal insurance design.

The rationale for this structure was clearly political. The majority of older adults have modest drug costs; in 2012, 54% of Medicare recipients, for example, had drug costs of less than \$2,000 per year.²⁵ The goal of the bill was, for a given federal budgetary cost, to deliver benefits to the largest number of elders regardless of their ultimate need for insurance coverage. This was accomplished by this upside-down drug benefit. Efforts to maximize votes by politicians, therefore, led to the use of low coinsurance for low expenditures and higher coinsurance for higher expenditures.

The donut hole was immediately criticized and was the source of much policy debate after the introduction of Part D. And for good reason: [Chandra et al. \(2021\)](#) find that when patients hit the donut hole, they cut back dramatically on medications of all types, leading to a sizeable rise in mortality. Due to a major feature of the ACA, the donut hole has now been “filled”; in 2021, the cost-sharing structure features 75% coverage until the “catastrophic limit,” then 95% coverage beyond that point. ■

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
reibaker@live.unc.edu

16.4 Controlling Costs in the Medicare Program

Unlike Medicaid, the focus of public policy debate on Medicare has not been about who should be eligible; there appears to be broad support for a program that universally covers older adults and people with disabilities. Given the broad consensus for universal coverage of older adults, the focus of debate has been on controlling the rapidly rising costs of this program. In its first 15 years of existence, the Medicare program grew exponentially, from \$64 million in federal expenditures in 1966 to \$32.1 billion in 1980. This rapid rise led policy makers to focus on controlling costs, through two different channels: the Prospective Payment System and Medicare Managed Care.

The Prospective Payment System

Like those in the private sector, the administrators of Medicare realized that its retrospective reimbursement of medical providers on the basis of their billed costs was a recipe for rapidly rising costs. This program was therefore a pioneer in moving toward prospective reimbursement, in which reimbursement is based on expected costs, not actual services delivered, in an effort to control overuse of medical care. In 1983, Medicare moved to a **Prospective Payment System (PPS)** for reimbursing hospitals. This system had several features:

1. All diagnoses for hospital admissions were grouped into 467 “Diagnosis Related Groups” or DRGs.
2. The government reimbursed hospitals a fixed amount based on the DRG of patient admission, regardless of the actual treatment costs of those patients. The reimbursement amounts were higher for more “severe” (higher-cost) DRGs.
3. The fixed amount of reimbursement was determined by a national standard for the cost of treating that DRG and a hospital-specific adjustment that more highly reimbursed teaching hospitals and those hospitals that treat many poor patients.

Prospective Payment System (PPS)

Medicare’s system for reimbursing hospitals based on nationally standardized payments for specific diagnoses.

Empirical Evidence on the Move to the PPS

In theory, PPS represented a classic prospective payment system, with incentives to treat patients as cost-efficiently as possible as hospitals are paid a fixed amount regardless of treatment intensity. Indeed, the effects of the PPS were striking. There was an enormous reduction in the treatment intensity of older adults within hospitals, a result consistent with the move from retrospective to prospective reimbursement incentives. The average length of a hospital stay for elderly patients fell from 9.7 days to 8.4 days in just one year, which was four times the rate of decrease over the previous two decades. In one Indiana hospital, the length of stay for hip fractures fell from almost 22 days to only 13 days. There was a 15% drop in admissions to intensive care units and a 16% drop in admissions to coronary (heart) care units.²⁶

Moreover, despite this enormous reduction in treatment intensity, there was no evidence of an adverse impact on patient outcomes. Mortality rates within one year of treatment were the same before and after this major policy change. This result is further evidence for the “flat of the curve” model shown in [Figure 15-8](#): as long as individuals are insured, treating them less intensively leads to little decline in their health.

The move to a PPS led to a sharp reduction in the rate of growth of hospital costs: after growing at 9.6% per year from 1967 to 1982, hospital costs under Medicare grew at only 3.0% per year from 1983 to 1988. Unfortunately, the PPS appears to lose its effect over time: from 1988 to 1997, hospital costs rose at a rate of 5.4% per year.

Problems with PPS

Why didn't the PPS solve the long-run cost growth problems of the Medicare program? Perhaps because it was not prospective enough. Almost immediately, the system ran into the problem of “DRG creep.” Medicare was paying a fixed price per diagnosis, but the choice of a diagnosis is something the hospital has some control over when patients are admitted (particularly for elderly patients, who often have many problematic conditions on admission). Thus, by labeling an admitted patient as having a more severe diagnosis, hospitals could change their DRG categorization to one for which the hospital would be reimbursed more highly. Indeed, there was a large increase in reported severity of admission diagnoses for older adults around the time PPS took effect!

Some examples were so egregious that they led to criminal prosecution of the largest hospital chain at the time in the United States, the Columbia/HCA corporation, for fraud. For example, in 1995 Columbia's Cedars Medical Center in Miami coded 93% of their Medicare cases with respiratory illness in the DRG for complex respiratory infection (with a reimbursement rate of \$5,700), and only 7% in the DRG for pneumonia with complications (with a rate of only \$1,700). Meanwhile, at Jackson Memorial, a non-Columbia-operated hospital located just across the street, only 28% of billings were in the complex respiratory infection category. Moreover, a pronounced shift in Cedars' billing pattern was shown to coincide exactly with its acquisition by the Columbia Corporation. In 1992, the hospital's last year of independent operation, only 31% of respiratory cases were billed at the highest rate; only one year later, after Columbia Corporation had bought the hospital, that number had risen to 76%. Ultimately, in 2000, HCA (as the company later renamed itself) pleaded guilty to 14 felonies and agreed to pay \$1.7 billion in civil and criminal penalties, the largest amount ever secured by federal prosecutors in a health care fraud case at the time.²⁷

This short-run problem has a longer-run manifestation, which is a problem with the design of the DRGs themselves. Almost half of the DRG designations are based not purely on diagnosis, but also on the actual treatment used for the patient. For example, someone entering the hospital with severe heart trouble might be given a diagnosis of "cardiac arrest, unexplained" (DRG 129), for which the hospital might be reimbursed \$5,000. Or they might receive one of many surgeries, including coronary bypass (DRG 106, with a reimbursement of \$33,000), pacemaker implantation (DRG 551, reimbursement of \$15,000), or in the most extreme case, a heart transplant (DRG 103, with a reimbursement of \$88,000).²⁸ With this categorization, there is effectively retrospective reimbursement: by performing a certain procedure, a provider can move the patient's case to a higher DRG and raise the reimbursement level.

Another problem with PPS is that it reimburses hospitals per admission, providing an incentive to raise hospital admissions. In particular, this has raised concerns about hospital re-admissions for the same diagnosis, whereby hospitals can make money by discharging patients before they are well and then re-admitting them for the same illness, but with new reimbursement. Indeed, a key feature of the ACA is penalties for hospitals that have high levels of re-admissions, and the introduction of these penalties was associated with a large drop in hospital re-admission rates.²⁹ But internal financial incentives from hospitals can have a pernicious effect: lawsuits filed against the for-profit hospital chain Health Management Associates

contended that the chain was paying doctors for admitting more Medicare patients to the hospital; for example, doctors were pressured to categorize any elderly patient with a temperature above the normal 98.7 degrees as having a “fever.”³⁰

Finally, another problem with the PPS has been that it applies only to one part of the medical system for treating older adults, but there is enormous substitutability across different parts of the medical system. One excellent example of this is provided by [Newhouse and Byrne \(1988\)](#), who studied rehabilitation hospitals, which are designed for long-term hospital stays that require little acute medical care (little constant monitoring and more long-term rehabilitation, such as for recovery from hip replacement surgery). Rehabilitation hospitals were originally exempt from the restrictions of the PPS. As a result, around the time of the PPS, there was an enormous shift of patients from acute-care hospitals to these exempt rehabilitation institutions, undoing a large part of the savings from the PPS change.

EMPIRICAL EVIDENCE

Short Stays in Long-Term Care Hospitals

As we've seen, an unintended effect of moving to PPS for hospitals was a shift toward more use of other types of facilities that were exempt from this system. Starting in 1997, Medicare tried to remedy these shifts by introducing PPS systems in these other facilities. But these alternatives can cause problems of their own.

An excellent example is provided by the study of [Kim et al. \(2015\)](#). They focused on long-term care hospitals, which they note “first emerged in the 1980s as a cost-effective alternative for patients who would otherwise require extended care in short-term acute care hospitals. Over time, long-term care hospitals have become the fastest growing and highest-paid providers of post-acute care [care delivered after hospitalization] in the Medicare program. From 1993 to 2011 the number of long-term care hospitals increased from 192 to 424, and Medicare expenditures on them soared from \$389 million to \$5.6 billion annually.”

Part of this growth was fueled by the implementation of PPS, which exempted these types of providers, creating incentives for acute-care hospitals to discharge patients quickly to these fee-for-service hospitals. To remedy this situation, in October 2002, the government introduced a PPS system for long-term care hospitals that was to be phased in over several years. To discourage the transfer of short-stay patients to these facilities, the system included a reduction in payments for patients discharged before a threshold length of stay (that varied by condition). This reduction was sizeable. For the case that the authors focus on, respiratory system diagnosis needing prolonged mechanical ventilation, the DRG reimbursement for 2015 was \$79,128. But if the patient left before their 29th day at the facility, that payment could be reduced to as low as \$31,376—a reduction of more than 50%.

The question addressed by this study is whether long-term care facilities responded to this incentive by keeping patients beyond the 29-day threshold. In particular, if hospitals were responding to this financial incentive, we should expect to see increased discharges right after the threshold when patients qualified for the higher payments.

patients qualified for the higher payments.

Of course, the problem with such an analysis is that there may be a natural reason patients with this diagnosis were discharged at 29 days—for example, maybe that is the typically recommended length of stay. To address this point, the authors used a comparison group of patients who were discharged in 2002 before the new system was introduced. For those patients, there was no particular financial incentive for discharge at the 29th day. Therefore, if there was no particular jump in discharge rates at the 29th day before 2002, but one emerged after 2002, then that suggests that financial incentives were playing an important role.

[Figure 16-2](#) suggests that, in fact, this was the case. This is a “hazard graph” of the type we used to analyze retirement in [Chapter 13](#): the x axis shows patient length of stay, while the y axis shows the percentage of patients leaving the long-term care hospital at that length of stay. The green line shows the distribution of lengths of stay in 2002, before the policy change, while the blue line shows the distribution from 2005–2010, after the policy was implemented. Before the policy change, there was a fairly constant discharge rate of 2% of patients at each length of stay until about 29 days, at which point it began to fall off. But after the policy change, there was a somewhat lower discharge rate at lengths of stay below 29 days (particularly at lengths of stay of 1 to 4 days), before spiking up sizably exactly at the 29-day threshold for receiving higher payments. Whereas 2% of patients were discharged at 29 days prior to the new policy, the share almost tripled after the policy change. Clearly, this change in reimbursement policy influenced hospital behavior.

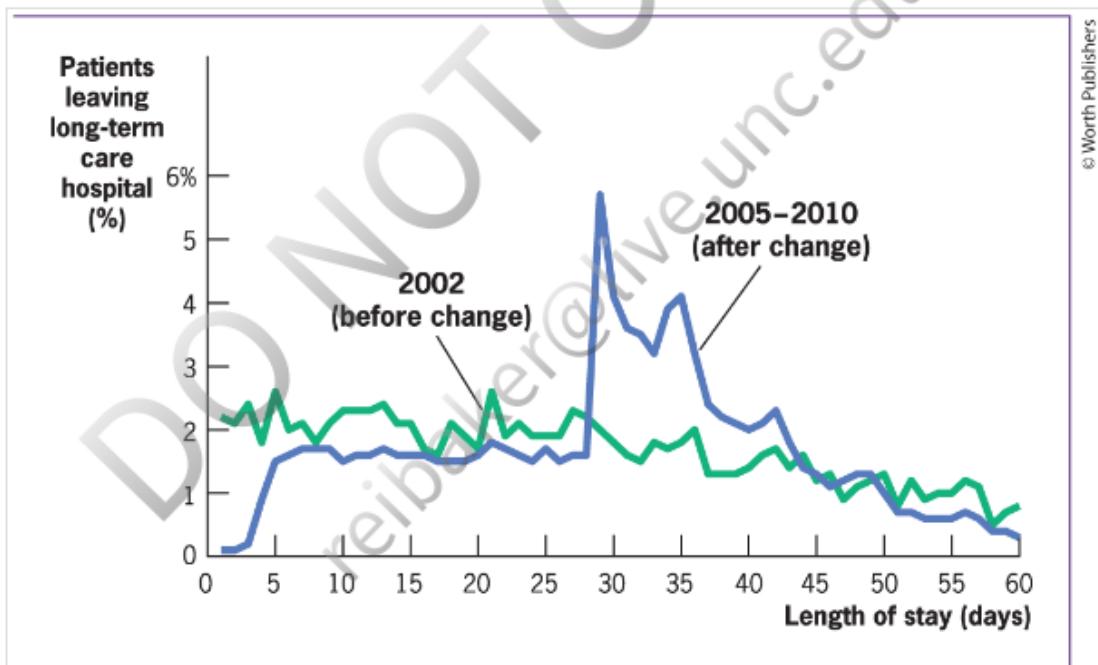


FIGURE 16-2 Timing of Hospital Discharges Before and After the Policy Change • The figure shows the percentage of patients discharged from long-term care hospitals at different lengths of stay. Before the policy change, 2% of patients were discharged at 29 days. After the change, the discharge rate almost tripled, hitting 5.7%.

Data from: [Kim et al. \(2015\)](#).



What can be done about such incentives? As discussed in the next section, the only way to truly incorporate such incentives is with a broad reimbursement system that features incentives for providers to coordinate care and internalize any inefficiency in how care is delivered—however, such a system is easier to describe than to implement. In the near term, unfortunately, the government is

System is easier to describe than to implement in the real world, unfortunately, the government is

moving in the other direction and Medicare introduced yet another reduction in reimbursement for very short stays at such facilities!

Lesson: The Difficulty of Partial Reform

These findings highlight a key problem with partial reform of provider reimbursement. If policy makers don't address systemwide incentives for overtreatment due to retrospective reimbursement, then partial solutions are like squeezing one corner of a pillow: the costs just move to the other corners. Pay hospitals based on diagnosis and patients suddenly appear sicker; reimburse one type of hospital more strictly and patients are moved elsewhere.

In 1997, recognition of this fact motivated the federal government to mandate that the prospective reimbursement system be applied to other sectors that were currently receiving retrospective reimbursement. Studies show that prospective reimbursement of home health care agencies and inpatient rehabilitation facilities lowered utilization significantly, with no adverse health outcomes.³¹

Just as with designing optimal insurance systems for workers, designing optimal reimbursement systems for providers reflects a trade-off. On the one hand, retrospective reimbursement systems do not provide sufficient incentives to control medical costs. On the other hand, a purely prospective system, which we have not yet achieved, might lead providers to cut care too much in order to make money. Thus, the optimal system would probably include some combination of both approaches. The existing evidence for the Medicare program suggests that it may err too much toward retaining key features of retrospective reimbursement.

Medicare Managed Care

The other avenue pursued by policy makers to control the costs of Medicare is to increase use of managed care in the Medicare program. Because managed care plans cost less per enrollee without obviously reducing the quality of care, the government could in theory save money by shifting enrollees to managed care.

Traditionally, all enrollees in Medicare received the same type of retrospectively reimbursed health insurance. Then, starting in 1985, the federal government allowed Medicare enrollees a choice of Medicare HMOs as well. These plans typically covered many of the out-of-pocket costs of Medicare so that enrollees were

less exposed to the program's copayments and deductibles. The plans also often provided other benefits not available through Medicare, most notably prescription drug coverage. A disadvantage for patients was that HMOs restricted their choice of provider and potentially engaged in other rationing devices to keep down costs that were not present in the traditional system. Despite this disadvantage, as [Figure 16-3](#) shows, enrollment in Medicare HMOs rose steadily to 16% of all enrollees in 1999, dipped somewhat, and then rose again after 2003 due to an increase in reimbursement rates to the managed care providers; enrollment growth has been particularly rapid in recent years.

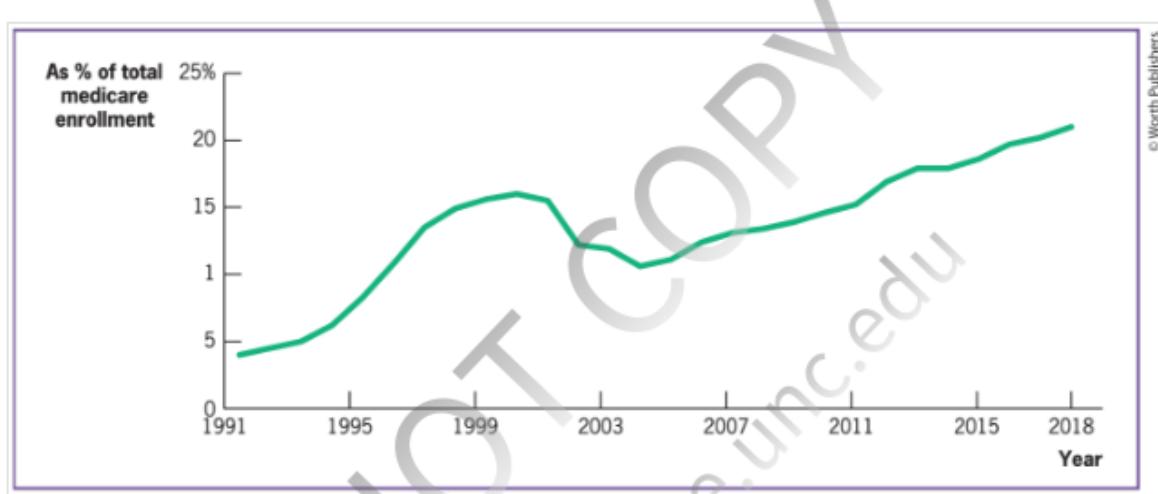


FIGURE 16-3 Managed Care Enrollment of Medicare Beneficiaries • When the managed care option was first introduced to Medicare, enrollment rose steadily, to a peak of 16% of Medicare beneficiaries by 1999. Enrollment then declined in the early 2000s because of the government's decision to lower reimbursement rates to managed care providers. An increase in these reimbursement rates in 2003 led to a subsequent further rise in managed care enrollment, which has continued since to surpass the 1999 peak.



The Medicare program endeavored to lower its costs by reimbursing HMOs only 95% of the average annual medical costs of enrollees who stayed in traditional Medicare (the “adjusted average per capita costs” or AAPCC). In this way, the thinking went, everyone won: the patients who chose Medicare HMOs got a package that they preferred, and the government saved money.

But was the government actually saving money? Recall from the previous chapter that there is strong evidence for positive selection into HMOs—only the healthiest patients choose this option. Such selection also operated strongly for older adults in Medicare. As a result, many Medicare HMOs had annual medical costs *for their enrollees* that were well below 95% of the AAPCC because they were enrolling only the healthiest people. Thus, most estimates suggest that the government was

actually losing money on the HMO option. How could the government lose money when it was paying only 95% of the average? Because the government was paying 95% of the average cost of the most expensive enrollees remaining on Medicare, while the HMO's true cost for its healthy enrollees was much lower. [Figure 16-4](#) illustrates the process of selection with a mythical Medicare program. Before HMOs are available, there are 300 individuals in this program. One-third of enrollees have average costs of \$1,000 per year, one-third have average costs of \$2,000 per year, and one-third have average costs of \$3,000 per year. So the average cost to the Medicare program is \$2,000 per enrollee, and the government spends a total of \$600,000.

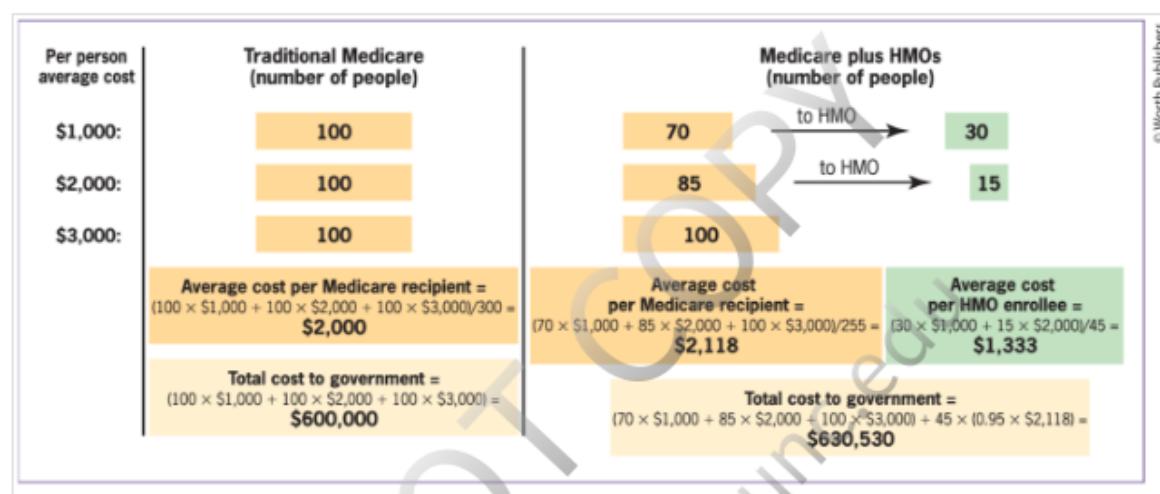


FIGURE 16-4 Incorporating HMOs into Medicare • Before HMOs are introduced (first column), there are 100 persons at each of three cost levels, with average Medicare costs of \$2,000 per recipient. When HMOs are introduced (second column), they are chosen primarily by the lowest-cost individuals, resulting in a higher average Medicare cost (\$2,118) and a low average HMO cost (\$1,333). Yet the government reimburses HMOs at 95% of the average Medicare cost (\$2,012), so the government spends \$30,530 more when the HMO option is allowed.



When HMOs are introduced, they are chosen by 30 of the individuals who cost \$1,000 per year, by 15 of the individuals who cost \$2,000 per year, and by none of the individuals who cost \$3,000 per year (because HMOs are preferred by the healthy). So the set of individuals in the HMOs cost on average \$1,333, while the set of individuals remaining in Medicare cost on average \$2,118. The government reimburses the HMOs 95% of the average cost of those on Medicare, or \$2,012 ($\$2,118 \times 0.95 = \$2,012$) per year. So the government is losing \$679 annually (the \$2,012 reimbursement minus the \$1,333 those individuals actually cost) on the average HMO enrollee. In other words, the government now has to pay for its remaining enrollees $(70 \times \$1,000) + (85 \times \$2,000) + (100 \times \$3,000) = \$540,000$, plus it pays the HMO \$2,012 for each of its 45 enrollees, or \$90,530, for a total of

\$630,530. So the government has lost more than \$30,000 by introducing this HMO option!

Recognizing this problem, in 1997 the government lowered further payments to HMOs. This led many HMOs, presumably the ones unable to attract the most healthy enrollees, to drop their Medicare lines of business. As a result, Medicare HMO enrollment, which peaked at 16% of enrollees in 1999, fell to 12.6% by 2003. Disturbed by this trend, Congress reversed itself as part of the Medicare Modernization Act of 2003 (better known for its introduction of prescription drug benefits for Medicare) and raised HMO reimbursement to at least 100% of the level of reimbursement for traditional Medicare (in 2004, it was 107%). This change had the effect of increasing managed care enrollment in Medicare, which rose above its previous peak to 22% of enrollees by 2009. It is unclear, however, why managed care enrollment per se should be a government goal; what is clear is that the government is once again losing money on reimbursing HMOs, relative to traditional Medicare!³² For this reason, reduced “overpayment” to Medicare managed care plans was a major source of financing for the new spending under the ACA; despite reduced Medicare Advantage reimbursement rates, as [Figure 16-3](#) shows, growth in use of Medicare Advantage has continued.

Should Medicare Move to a Full-Choice Plan? Premium Support³³

The problem with the current approach to choice in Medicare is that the government must make some estimate of how much to reimburse managed care plans, and that estimate may be too high (costing the government money) or too low (leading the HMOs to exit the program). Another alternative is to move to a system of full choice among health plans, leaving older adults with the type of decision faced by employees at firms that offer multiple health insurance plans. In this type of system, commonly known as **premium support**, elders would receive a voucher for a certain amount. They could then choose from a range of options, including the traditional Medicare program. If they choose an option that costs less than their voucher amount, they keep the difference; if they choose an option that costs more, they pay the difference. As we discuss in the application, such a system would introduce competition into Medicare and allow the government to more easily set a reimbursement level, but at the cost of redistributing from the sickest enrollees to the most healthy.

premium support

A system of full choice among health care plans for Medicare enrollees, whereby they receive a voucher for a certain amount that they can apply to a range of health insurance options (either paying or receiving the difference between plan premiums and the voucher amount).

APPLICATION

A Premium Support System for Medicare



A typical premium support plan is illustrated in the top panel of [Table 16-3](#). An elderly person would have a choice in their area of, for example, three plans, A, B, and C. The costs of those plans per year are \$1,800, \$2,000, and \$2,500, respectively. The government could set its voucher at any level, but suppose the government chooses to reimburse the cost of the median-cost plan in the area (\$2,000 in this case). Then individuals who choose plan A would get a rebate of \$200. Individuals who choose plan B would get nothing, and individuals who choose plan C would have to pay \$500.

TABLE 16-3 Premium Support Systems with and without Adverse Selection

Full-choice Medicare (before adverse selection)			
Plan	Plan cost (per person)	Voucher (median plan cost)	Individual payment
A	\$1,800	\$2,000	-\$200
B	\$2,000	\$2,000	0
C	\$2,500	\$2,000	\$500

Full-choice Medicare (after adverse selection)			
Plan	Plan cost (per person)	Voucher (median plan cost)	Individual payment
A	\$1,600	\$2,100	-\$500
B	\$2,100	\$2,100	0
C	\$3,000	\$2,100	\$900

Under a typical premium support system, individuals pay the difference between the cost of their plan and the median-cost plan, as shown in the top panel. Adverse selection, however, will cause sicker patients to choose the most expensive plans, making these plans even more expensive, while the least costly plans fall in price as they are chosen by healthier individuals. In the long term, a voucher system thus ends up rewarding the healthy and costing the sick more (bottom panel).

The advantages of a premium support system mirror the advantages of voucher systems for education (discussed in [Chapter 11](#)). First, such a system respects consumer sovereignty by allowing individuals to choose the health plan that best matches their taste rather than forcing them into one government-provided option. Second, such a system promotes efficiency in medical care delivery by allowing individuals to shop across plans. Both HMOs and Medicare will have to produce care as efficiently as possible so that they can offer the lowest possible premiums to attract enrollees. There is less incentive to do so under today's Medicare program because enrollees pay the same premium

amount regardless of which option they choose.

Finally, such a system solves the problem of “appropriate” reimbursement levels for managed care plans by simply letting the market work. For example, the government would simply announce that it would reimburse the amount of the median-cost plan in the area. Plans would then compete to have low costs in order to attract enrollees, and after this competition, the government could simply find the median price and set that as its reimbursement level. There would be no more guessing at the right reimbursement level that fits all HMOs.

Offsetting these positive attributes of health care choice are two clear concerns. The first is adverse selection. If the government simply reimburses a flat amount and makes individuals pay on the margin for more expensive plans (or allows them to pocket savings on less expensive plans), then we run into the same problem that we first discussed in the context of Harvard University in [Chapter 12](#) ([pp. 331–332](#)): healthy individuals choose the less expensive plans, raising costs even further for sicker individuals, who prefer the more generous plans.

The bottom panel of [Table 16-3](#) illustrates what the premium support system might look like after adverse selection has taken its toll. The healthiest individuals have moved into plan A, lowering its costs further to \$1,600 per year. Some of these individuals came from plan B, so plan B now has sicker enrollees on average, and its costs have risen to \$2,100. A few individuals from plan C might have moved to plan A, and more from plan C have moved to plan B, so plan C now has just the sickest enrollees, and its costs have risen to \$3,000. Now, with the voucher amount tied to the median-cost plan, those in plan A get a rebate of \$500 per year, while those in plan C must pay \$900 per year.

In a system that reimburses the cost of the median health plan offered, the adverse selection problem could lead to very large rebates for those who require little health care (who generally choose plan A) and very large costs for those who need a lot of it (who generally choose plan C). At its heart, therefore, this adverse selection problem is primarily an income distribution issue: Are we willing to permit healthier elderly people to save money and sicker elderly people to pay more in order to introduce competition into the Medicare system? The appropriate level of plan choice in Medicare reflects the trade-off between the benefits of competition and the costs of redistribution from the sick to the healthy.³⁴

The second major concern is poor decision making. In the typical economics model, more choice can only make individuals better off, but there is a growing body of literature in behavioral economics that says additional choices can worsen the outcome of the decision process. Studies in a wide variety of contexts show that more choices lead to less participation in a market and in less satisfaction with the choices that are made. The literature on Medicare Part D reviewed in [Abaluck and Gruber \(2016\)](#) shows that Medicare recipients facing the large choice sets of prescription drug plans (typically about 50 options) are doing a poor job in making their choices and are not learning how to choose better over time. And follow-up studies in the broader health insurance context confirm these “choice inconsistencies”: most striking is evidence that shows that individuals frequently choose “dominated” plans in their health insurance choice sets, that is, plans that are worse along every possible dimension than other available alternatives.³⁵

In principle, both of these problems can be solved with technical fixes. Redistribution due to adverse selection could be offset by appropriate risk adjustment, whereby insurers who enroll the sickest patients receive transfers from insurers who enroll the healthiest ones. There have been enormous advances in risk adjustment in recent years, and very sophisticated approaches to adjustment are now used by the government and private insurers. Even these sophisticated adjustors are very imperfect, however, so it is impossible to offset fully the health risk differences across plans.³⁶ Likewise, poor decision making could be addressed with proper “decision support” tools that

Likewise, poor decision making could be addressed with proper “decision support tools that carefully guide customers through the complicated aspects of their insurance plan decision. Such tools are still in their infancy, however, with little evidence to date that they have dramatically improved insurance decision making by individuals. ■

Gaps in Medicare Coverage

While most of the policy debate around Medicare has focused on cost control, another important source of debate has been whether and how to enrich the Medicare benefits package. Medicare insurance is significantly less generous than most private insurance because of its high copayments and deductibles and because it leaves some medical goods and services uncovered. Individuals fill these coverage gaps in Medicare in one of three ways:

1. Low-income elderly individuals are entitled to more generous coverage under the Medicaid program or through subsidies to private prescription drug plans.
2. About 28% of Medicare beneficiaries also have supplemental employer-sponsored insurance, which fills in many of these gaps.
3. Many retirees not covered by Medicaid or their own retiree health insurance buy individual “Medi-gap” policies from insurance companies that fill these gaps.

An important problem with these three ways of filling the gaps in Medicare coverage is that they exert a *negative financial externality* on the Medicare program. As discussed earlier, patient coinsurance for medical care costs reduces the total amount of medical care used by Medicare enrollees. Thus, when other forms of insurance cover Medicare’s deductibles and coinsurance, the amount of medical care used increases. Because the costs of most of this medical care (what is spent beyond copayments) is covered by Medicare, there is a negative externality imposed on Medicare by these other forms of insurance: those holding this supplementary insurance raise Medicare costs without bearing the costs of doing so. If I have to pay \$20 for my \$100 doctor’s visit under the traditional Medicare plan, I may not go to the doctor, so Medicare incurs no costs; but if my Medi-gap policy covers that \$20, then I will go, and the Medicare program will see its costs increase by \$80 through no action of its own. The availability of Medi-gap has caused me to spend \$80 more of Medicare’s money, but my Medi-gap insurer and I bear none of this \$80 in increased public sector costs. [Cabral et al. \(2018\)](#) show that having Medi-gap increases an individual’s Medicaid spending by 22%.

An important problem with these three ways of filling the gaps in Medicare coverage is that they exert a *negative financial externality* on the Medicare program. As discussed earlier, patient coinsurance for medical care costs reduces the total amount of medical care used by Medicare enrollees. Thus, when other forms of insurance cover Medicare's deductibles and coinsurance, the amount of medical care used increases. Because the costs of most of this medical care (what is spent beyond copayments) is covered by Medicare, there is a negative externality imposed on Medicare by these other forms of insurance: those holding this supplementary insurance raise Medicare costs without bearing the costs of doing so. If I have to pay \$20 for my \$100 doctor's visit under the traditional Medicare plan, I may not go to the doctor, so Medicare incurs no costs; but if my Medi-gap policy covers that \$20, then I will go, and the Medicare program will see its costs increase by \$80 through no action of its own. The availability of Medi-gap has caused me to spend \$80 more of Medicare's money, but my Medi-gap insurer and I bear none of this \$80 in increased public sector costs. [Cabral et al. \(2018\)](#) show that having Medi-gap increases an individual's Medicaid spending by 22%.

16.5 Long-Term Care

The discussion of health care thus far in this chapter has focused on acute medical care, such as for flu or a heart attack. A growing share of medical spending, however, is devoted to chronic long-term care needs, such as nursing home stays for older adults and people with disabilities. In 1960, only 3.4% of health care dollars were spent on **long-term care**. In 2019, 8% of health care spending, or \$286.2 billion per year, was on long-term care.³⁷

long-term care

Health care delivered to the disabled and elderly for their long-term rather than acute needs, either in an institutional setting (a nursing home) or in their homes.

This care is delivered primarily in two forms:

1. *Institutional care* provided in nursing homes. Reimbursement of the cost of such institutions (where infirm individuals live full-time) accounts for 60% of long-term care costs. Medicaid is the primary insurer of long-term nursing home stays, covering 62% of all residents, but due to relatively low reimbursement this amounts to only 29% of total costs. Medicare is the primary insurer of short-term nursing home stays, covering the first 100 days in the home for those who transition from hospitals with higher reimbursement rates, and covers 22% of costs nationally.³⁸
2. *Home health care*, where nurses and other aides provide care in the patient's home, accounts for the remaining 40% of long-term care costs. Medicare is the primary insurance payer for home health care, covering 39% of such costs nationally through its Part A home health care benefit, although Medicaid also contributes coverage of nearly a third of costs.³⁹ Since 1980, there has been an enormous shift from institutional to home health care.

Financing Long-Term Care

The major debate over long-term care is about financing. Currently, nursing home costs are financed mostly by private payers (individual self-insurance and, to a small but growing extent, private long-term care insurance) and by Medicaid. Individuals who enter nursing homes begin by paying costs out of their own savings; usually those savings are rapidly drawn down by nursing home stays, which cost more than \$93,000 per year.⁴⁰ When savings are drawn below a threshold level, individuals qualify for state programs that pick up the cost of nursing homes under Medicaid. Individuals are therefore insured against nursing home costs, but only if

they use up all their personal savings first. In this sense, Medicaid imposes an implicit tax on assets: Medicaid will provide financing only once assets are low, so having a lot of assets implies forgoing the right to a government-financed nursing home stay.

This financing system has several problems. First, those who wish to leave money behind when they die have no protection against losing their entire estate to nursing home costs. Second, individuals have an incentive to cheat by hiding their wealth in forms that cannot be found or accessed by Medicaid authorities, thereby more quickly qualifying for public insurance while preserving their wealth. Finally, this system *crowds out* savings for old age because people can qualify for Medicaid only once they have spent their wealth. This situation parallels our discussion of education in [Chapter 11](#): it leads those who would like to save for a somewhat high-quality nursing home when they are old to not save at all instead so they can take advantage of the free Medicaid entitlement. There is a private long-term care insurance market that, in principle, could solve these difficulties. Despite rapid growth in the past decade, however, this market remains small, paying only about 8.1% of nursing home costs.⁴¹ There are many potential causes of failure in this market, most notably adverse selection based on private information about the likely risk of nursing home stays. Thus, there may be an argument for replacing the government's current patchwork of long-term care financing with a social insurance system. As always, the potential cost of such a move would be increased by moral hazard, in the form of increased and long-term use of nursing homes by older adults if the costs of such care are insured, although existing research suggests relatively little moral hazard in nursing home use.⁴² There was an attempt to provide such a program as part of the ACA, labeled the "CLASS Act," but it was not implemented based on concerns about its financial sustainability.

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
reibaker@live.unc.edu

This financing system has several problems. First, those who wish to leave money behind when they die have no protection against losing their entire estate to nursing home costs. Second, individuals have an incentive to cheat by hiding their wealth in forms that cannot be found or accessed by Medicaid authorities, thereby more quickly qualifying for public insurance while preserving their wealth. Finally, this system *crowds out* savings for old age because people can qualify for Medicaid only once they have spent their wealth. This situation parallels our discussion of education in [Chapter 11](#): it leads those who would like to save for a somewhat high-quality nursing home when they are old to not save at all instead so they can take advantage of the free Medicaid entitlement. There is a private long-term care insurance market that, in principle, could solve these difficulties. Despite rapid growth in the past decade, however, this market remains small, paying only about 8.1% of nursing home costs.⁴¹ There are many potential causes of failure in this market, most notably adverse selection based on private information about the likely risk of nursing home stays. Thus, there may be an argument for replacing the government's current patchwork of long-term care financing with a social insurance system. As always, the potential cost of such a move would be increased by moral hazard, in the form of increased and long-term use of nursing homes by older adults if the costs of such care are insured, although existing research suggests relatively little moral hazard in nursing home use.⁴² There was an attempt to provide such a program as part of the ACA, labeled the "CLASS Act," but it was not implemented based on concerns about its financial sustainability.

16.6 Health Care Reform and the ACA

The United States has struggled with fundamental health care reform for almost a century, since early reform efforts in 1915. As reviewed in [Starr \(2011\)](#), the ACA was viewed as a culmination of that struggle, but continued controversy over the program leaves its ongoing status uncertain.

The Historical Impasse

Health care reformers in the United States have always struggled to find a compromise between two extremes. On the one side were advocates for a public single-payer health care system such as the one that exists in Canada. Under such a system, the government insures all citizens by putting them into a single publicly run health insurance plan. Costs are controlled through an explicit nationwide health care budget that imposes true prospective reimbursement on the health care system as a whole. For example, the government establishes a network of providers (physicians, hospitals, and so on) in each region and pays that network a fixed amount to cover all the medical costs for every person in that region. This amount could then be increased slowly to limit medical care spending growth.

Such a plan has major advantages. First, it fully solves the problem of the uninsured. Second, it reduces the administrative costs inherent in the U.S. medical care system: the administrative costs of Canada's national health insurance system are only 2.8%, compared to the 12% costs on average for private health insurance in the United States.⁴³ Third, controlling costs through a single government payer can help address the problems with piecemeal cost controls discussed earlier this chapter. Finally, it would solve the many inequities and inefficiencies arising from the patchwork of health insurance coverage available in the United States; for example, it would resolve the problem of job lock (discussed in the previous chapter) as individuals would be able to switch jobs without worrying about losing their employer-based insurance.

Public national health insurance also has major disadvantages, however. First, it would require massive new government expenditures because the government would now be paying the insurance costs of every citizen. A large part of this cost would be offset by reduced spending on existing public insurance for the non-elderly, but most would require new public-sector revenues.

► Quick Hint

One confusing aspect of discussions of public health insurance is that while public expenditures would rise dramatically, there would be an almost equally large (or even larger) reduction in private insurance expenditures. Thus, the rise in total social costs of health care would be small (or even negative) compared with the actual costs to the government. In theory, it is these social costs that matter, not the budgetary expenditures of the government. In practice, however, the expenditures by the government matter for two reasons. First, as discussed at more length in the tax chapters, there may be a deadweight loss arising from the need to increase government revenues to pay for these expenditures that does not arise from private spending (whether that deadweight loss is larger than the administrative gains from public insurance is unclear). Second, and more important in practice, moving from private financing of health insurance through employer expenditures to public financing is like moving from a hidden tax to an explicit tax. While economists can talk until they are blue in the face about how this is just shifting from one payer to another, the typical voter sees it as a massive tax increase and will be reluctant to support it.

Second, nationwide budgeting as a means of controlling health care costs is a very blunt instrument that may not allow doctors to use a technology that is worth its high cost. For example, suppose that the U.S. government had imposed a global budget on health care spending of 5% of gross domestic product (GDP) in 1950. While the health care sector would likely be much more efficient today, it is also likely that many beneficial medical advances would not be available. Once again, it seems unlikely that many of us would return to the world of 1950 medical care, even at 1950s prices, if we had the choice. It is possible that we could be saying the same thing in 2070 about not wanting to return to 2020 care at 2020 prices!

Finally, national health insurance faces political hurdles that are likely insurmountable in the near term. The health insurance industry in the United States has revenues of nearly \$960 billion per year, making it a very concentrated interest that would provide a powerful roadblock to this approach.⁴⁴ Moreover, the majority of Americans, particularly those working for large firms with a choice of plans, are quite content with their private health insurance. Telling them that they have to give up that insurance so that a minority of Americans can get covered will be a very difficult political sell (as witnessed by the attacks on health care reforms both under President Clinton and President Obama that were based on fear of restricting insurance choice).

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DO NOT COPY
reibaker@live.unc.edu

More significantly, in his campaign President Biden proposed a “public option,” a concept developed by [Hacker \(2009\)](#) that was included in early drafts of the ACA. The public option is designed to split the difference between those who would advocate for a single government insurer and those who believe in insurer competition, by offering a public insurance plan (such as Medicare) as an option on the insurance exchanges. Such a plan would have three advantages. First, it would increase competition on the exchanges; [Dafny, Gruber, and Ody \(2015\)](#) estimate that each additional insurer added to the exchange significantly lowers premiums. Second, it would provide an avenue to enroll low income individuals in states that have not yet expanded Medicaid; in those states, the public option could be offered to those below the poverty line who are currently without insurance options under the ACA.

Third, the public option would potentially provide a much lower cost insurance option for those who buy insurance, since public programs like Medicare pay lower regulated prices for health care services. It is this feature that makes the public option particularly controversial. Private insurers argue that this gives the public option an unfair advantage that would tilt the playing field away from private plans; one estimate suggests that a public option would reduce private enrollment by 10 million.^{[62](#)}

As of this writing, it is uncertain which of the Biden Administration proposals will become reality. But what seems clear at this point is that the ACA is here to stay, with continued fighting back and forth about its reach and costs.

EMPIRICAL EVIDENCE

The Impact of the ACA on Mortality

A key question about the ACA from its passage has been: will it improve the health of Americans? And there is perhaps no better indicator of the effect of a public intervention on health than its effects on mortality. Fortunately, however, mortality is a relatively rare outcome in the United States for all but the oldest old. This raises a problem of statistical “power”: with rare outcomes, it is hard to derive a statistically meaningful conclusion without large samples.^{[63](#)} Two recent studies, however, have managed to crack the power conundrum, and they have delivered strong evidence that the ACA has lowered mortality in the United States.

The first of these studies is [Miller et al. \(2019\)](#). They take a standard quasi-experimental approach to analyzing the effects of the Medicaid expansions on the mortality of older Americans (those age 55–64), comparing the mortality outcomes over time in states that did and did not expand their Medicaid programs under the ACA. They are able to use administrative death records so that they can study mortality among a sample of almost 700,000 people. While they follow the same approach laid out in

the previous empirical application on Medicaid expansions, they extend to the newer standard for empirical work, which is to complement regression analysis with an “event study” illustration.

The results of their analysis are shown in [Figure 16-6](#). The x axis shows time relative to the date that Medicaid expands in a given state: for example, time -2 is two years before the expansion, time 0 is the year of the expansion, and time 2 is two years after expansion. All of the expansions are normalized in this way; if a state expands Medicaid in 2014, year 0 is 2014; if the expansion is in 2016, year 0 is 2016.

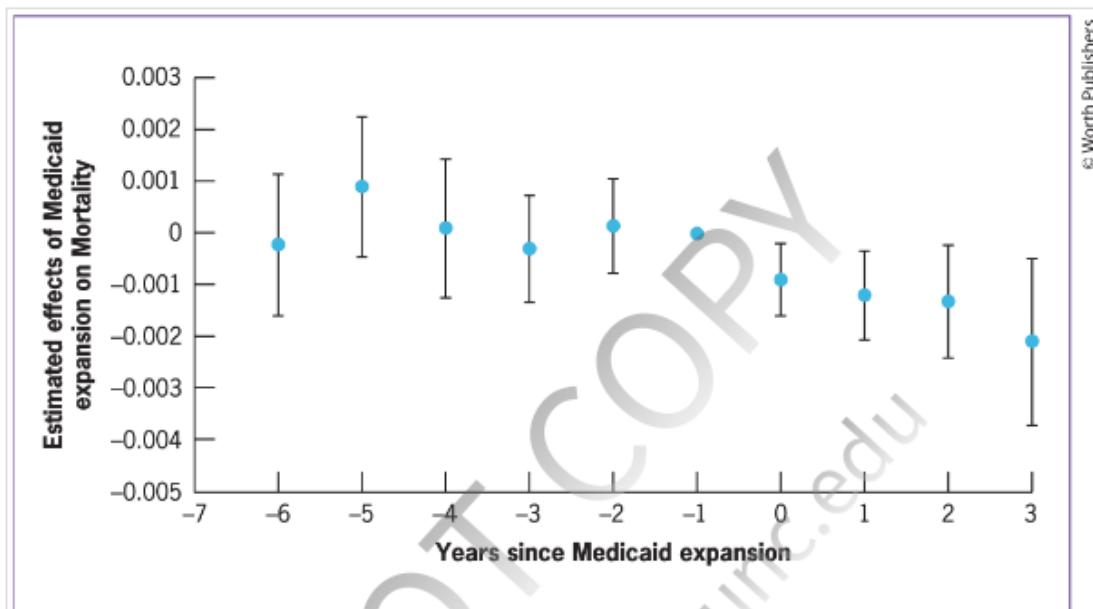


FIGURE 16-6 Decrease in Mortality Following Medicaid Expansion • The blue dots show the coefficient estimate for the effects of Medicaid expansion on mortality, and the black bars show confidence intervals. The negative numbers on the Y axis represent the years before Medicaid is expanded. These data show a downward trend in mortality following the expansion of Medicaid.

Data from: [Miller, Johnson, and Wherry \(2020\)](#).



The y axis shows the magnitude of the estimated effects of expansion on mortality. The dots show the coefficient estimate, while the bars show the “confidence intervals” of the estimate, which reflects the precision of the estimated coefficient.⁶⁴ That is, any estimate where the bars cross zero at any point is not statistically significant. The estimate for year -1 is set to zero by definition, as all other estimates are expressed relative to that year.

The results show that in every year before Medicaid expands, there is no significant impact on mortality. The estimates are always around zero and in no cases do the confidence intervals not cross the horizontal line at zero. This is comforting for the reasons discussed in the earlier empirical evidence box: it shows that there were no underlying trends in mortality that might have been biasing the estimates (which would show up as effects even before the law passed). After passage of Medicaid expansion, mortality falls significantly—and the impact appears to grow over time (although, as the wider bars show, it becomes less precise over time as well). The authors estimate that the Medicaid expansions lowered mortality in this age group by 9.4%.

An even more innovative study is that of [Goldin et al. \(2019\)](#), who created an amazing field experiment. As part of an effort to increase compliance with the ACA individual mandate, the Internal Revenue Service (IRS) sent letters in late 2016 and January 2017, to individuals who had paid the

individual mandate penalty on their 2015 taxes. The letters “informed recipients that they had paid a penalty in 2015; provided information about the penalty and plan costs for 2017; and provided instructions for recipients to investigate the availability of Exchange and Medicaid coverage through [healthcare.gov](#).” The IRS sent 3.9 million such letters.

At the same time, the IRS created a randomized control group of 1.2 million taxpayers who had also paid the individual mandate penalty in 2015 but who did not get a reminder letter. The randomization across those who did and did not get the letter allowed researchers to assess the impact of receiving a letter on both insurance coverage—and mortality.

The study first explored the impact of receiving the letter on having insurance coverage, as measured by complying with the mandate on subsequent tax filings. The authors find that receiving a letter, relative to the control group, significantly increased the odds of having insurance coverage—that is, being reminded about the mandate penalty increased compliance with the mandate. They estimate that there was a 1.3% increase in the odds of having insurance for the treatment group relative to controls.

While highly statistically significant, this is a small effect. Fortunately, for studying mortality effects, they had access to a statistical match of the experimental sample to government death records, allowing them to assess mortality effect for millions of taxpayers. For the sample of 45–64 year olds that they focus on, in fact they find that receipt of the letter did significantly reduce mortality, by about 0.06%.

To interpret this small effect, they express it in terms of mortality per person newly insured. That is, they find that letter receipt increases insurance coverage, and that it also lowers mortality; dividing the latter by the former, they find that for each person gaining coverage mortality falls by over 2%.

The calculated mortality per person insured is higher than that found by Miller et al., who estimate that mortality per person insured fell within the range of 0.8 to 1.03%. This highlights the important point that multiple empirical studies on the same topic can give convincing yet different answers. But the bottom line is clear: the ACA has had a positive impact on the health of Americans.

Printed by: reibaker@live.unc.edu. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

DONOTCOPY
reibaker@live.unc.edu

More significantly, in his campaign President Biden proposed a “public option,” a concept developed by [Hacker \(2009\)](#) that was included in early drafts of the ACA. The public option is designed to split the difference between those who would advocate for a single government insurer and those who believe in insurer competition, by offering a public insurance plan (such as Medicare) as an option on the insurance exchanges. Such a plan would have three advantages. First, it would increase competition on the exchanges; [Dafny, Gruber, and Ody \(2015\)](#) estimate that each additional insurer added to the exchange significantly lowers premiums. Second, it would provide an avenue to enroll low income individuals in states that have not yet expanded Medicaid; in those states, the public option could be offered to those below the poverty line who are currently without insurance options under the ACA.

Third, the public option would potentially provide a much lower cost insurance option for those who buy insurance, since public programs like Medicare pay lower regulated prices for health care services. It is this feature that makes the public option particularly controversial. Private insurers argue that this gives the public option an unfair advantage that would tilt the playing field away from private plans; one estimate suggests that a public option would reduce private enrollment by 10 million.⁶²

As of this writing, it is uncertain which of the Biden Administration proposals will become reality. But what seems clear at this point is that the ACA is here to stay, with continued fighting back and forth about its reach and costs.

EMPIRICAL EVIDENCE

The Impact of the ACA on Mortality

A key question about the ACA from its passage has been: will it improve the health of Americans? And there is perhaps no better indicator of the effect of a public intervention on health than its effects on mortality. Fortunately, however, mortality is a relatively rare outcome in the United States for all but the oldest old. This raises a problem of statistical “power”: with rare outcomes, it is hard to derive a statistically meaningful conclusion without large samples.⁶³ Two recent studies, however, have managed to crack the power conundrum, and they have delivered strong evidence that the ACA has lowered mortality in the United States.

The first of these studies is [Miller et al. \(2019\)](#). They take a standard quasi-experimental approach to analyzing the effects of the Medicaid expansions on the mortality of older Americans (those age 55–64), comparing the mortality outcomes over time in states that did and did not expand their Medicaid programs under the ACA. They are able to use administrative death records so that they can study mortality among a sample of almost 700,000 people. While they follow the same approach laid out in

the previous empirical application on Medicaid expansions, they extend to the newer standard for empirical work, which is to complement regression analysis with an “event study” illustration.

The results of their analysis are shown in [Figure 16-6](#). The x axis shows time relative to the date that Medicaid expands in a given state: for example, time -2 is two years before the expansion, time 0 is the year of the expansion, and time 2 is two years after expansion. All of the expansions are normalized in this way; if a state expands Medicaid in 2014, year 0 is 2014; if the expansion is in 2016, year 0 is 2016.

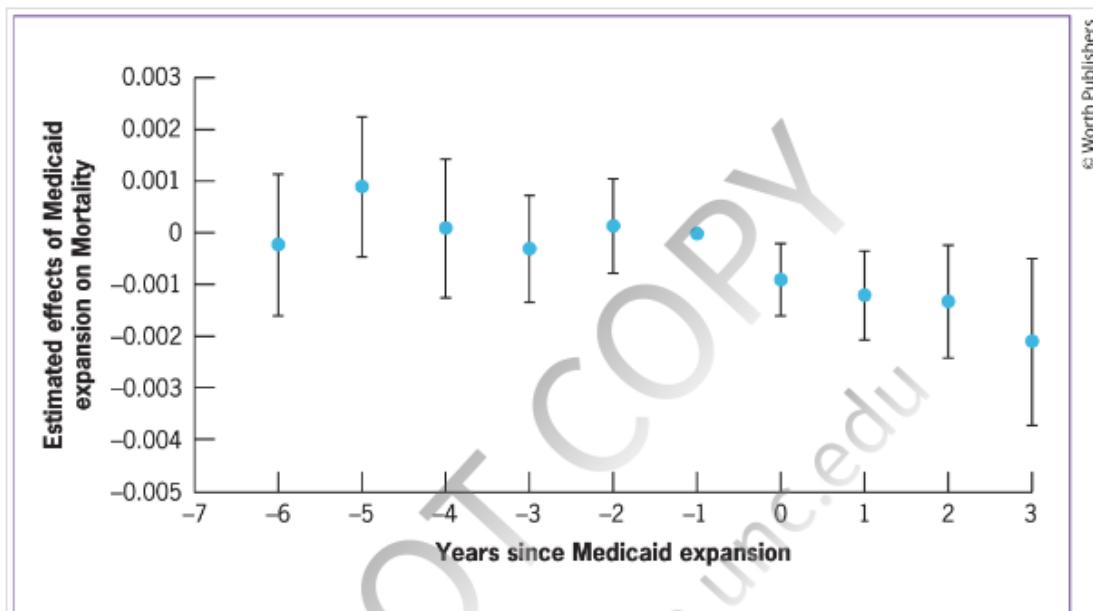


FIGURE 16-6 Decrease in Mortality Following Medicaid Expansion • The blue dots show the coefficient estimate for the effects of Medicaid expansion on mortality, and the black bars show confidence intervals. The negative numbers on the Y axis represent the years before Medicaid is expanded. These data show a downward trend in mortality following the expansion of Medicaid.

Data from: [Miller, Johnson, and Wherry \(2020\)](#).



The y axis shows the magnitude of the estimated effects of expansion on mortality. The dots show the coefficient estimate, while the bars show the “confidence intervals” of the estimate, which reflects the precision of the estimated coefficient.⁶⁴ That is, any estimate where the bars cross zero at any point is not statistically significant. The estimate for year -1 is set to zero by definition, as all other estimates are expressed relative to that year.

The results show that in every year before Medicaid expands, there is no significant impact on mortality. The estimates are always around zero and in no cases do the confidence intervals not cross the horizontal line at zero. This is comforting for the reasons discussed in the earlier empirical evidence box: it shows that there were no underlying trends in mortality that might have been biasing the estimates (which would show up as effects even before the law passed). After passage of Medicaid expansion, mortality falls significantly—and the impact appears to grow over time (although, as the wider bars show, it becomes less precise over time as well). The authors estimate that the Medicaid expansions lowered mortality in this age group by 9.4%.

An even more innovative study is that of [Goldin et al. \(2019\)](#), who created an amazing field experiment. As part of an effort to increase compliance with the ACA individual mandate, the Internal Revenue Service (IRS) sent letters in late 2016 and January 2017, to individuals who had paid the

individual mandate penalty on their 2015 taxes. The letters “informed recipients that they had paid a penalty in 2015; provided information about the penalty and plan costs for 2017; and provided instructions for recipients to investigate the availability of Exchange and Medicaid coverage through [healthcare.gov](#).” The IRS sent 3.9 million such letters.

At the same time, the IRS created a randomized control group of 1.2 million taxpayers who had also paid the individual mandate penalty in 2015 but who did not get a reminder letter. The randomization across those who did and did not get the letter allowed researchers to assess the impact of receiving a letter on both insurance coverage—and mortality.

The study first explored the impact of receiving the letter on having insurance coverage, as measured by complying with the mandate on subsequent tax filings. The authors find that receiving a letter, relative to the control group, significantly increased the odds of having insurance coverage—that is, being reminded about the mandate penalty increased compliance with the mandate. They estimate that there was a 1.3% increase in the odds of having insurance for the treatment group relative to controls.

While highly statistically significant, this is a small effect. Fortunately, for studying mortality effects, they had access to a statistical match of the experimental sample to government death records, allowing them to assess mortality effect for millions of taxpayers. For the sample of 45–64 year olds that they focus on, in fact they find that receipt of the letter did significantly reduce mortality, by about 0.06%.

To interpret this small effect, they express it in terms of mortality per person newly insured. That is, they find that letter receipt increases insurance coverage, and that it also lowers mortality; dividing the latter by the former, they find that for each person gaining coverage mortality falls by over 2%.

The calculated mortality per person insured is higher than that found by Miller et al., who estimate that mortality per person insured fell within the range of 0.8 to 1.03%. This highlights the important point that multiple empirical studies on the same topic can give convincing yet different answers. But the bottom line is clear: the ACA has had a positive impact on the health of Americans.

16.7 Conclusion

The past decade has seen the most fundamental transformation of the U.S. health care system since the 1960s. The ACA held the potential to address many shortcomings in our health insurance system, and it greatly reduced the ranks of the uninsured. Early evidence suggests that the ACA significantly improved health and may have been a contributor to a historic slowdown in health care costs. The law's initial failure to achieve popularity, however, led to significant scaling back under the Trump Administration. But the law largely remains in place, with the current administration proposing to expand it further.

This is just the beginning of a series of reforms that will be needed to address the larger, long-run problem of controlling health care costs. Attempts at cost control have to recognize that the U.S. government already contributes about half of the health care spending in the United States through its Medicare and Medicaid programs. Therefore, important changes in those programs will be required to slow health care cost growth.

HIGHLIGHTS

- The Medicaid program serves low-income families, low-income people with disabilities, and low-income older adults.
- Low-income families are eligible for Medicaid if they are on cash welfare; in most states, low-income children and pregnant people are eligible if they have incomes below 200% of the poverty line.
- The Medicaid program has been shown to reduce the number of uninsured people (although there is also significant crowd-out of private health insurance coverage), increase health care utilization, and improve health in a cost-effective manner.
- The Medicare program serves older adults and people with disabilities, providing universal coverage with significant enrollee copayments.
- The Medicare program has tried to limit its cost growth through restrictions on provider reimbursement, most notably with a move to prospective reimbursement of hospitals (PPS), which has lowered costs without significantly reducing quality of care.
- The Medicare program has also tried to introduce choice across insurance options with limited success. A premium support plan, which would more aggressively induce choice in Medicare, could control costs but would have adverse effects on the living standards of the sickest beneficiaries.
- A large and rising share of medical spending is on long-term care, and there is a major debate over appropriate division of the financing of these costs between individuals and the government.
- The century-long debate over fundamental health care reform in the United States was addressed in many important ways by the ACA. This law builds on a successful reform in Massachusetts to attempt to cover most of the uninsured and repair the market failures in the nongroup insurance market. The law was initially successful in increasing insurance coverage, but recent legislative and regulatory actions have weakened it.
- The longer-run problem still facing the United States, however, is the rapid rise in health care costs. This is a difficult problem to resolve because much of past health care spending growth has resulted in improved health outcomes for the American population. Future reforms will face the challenge of slowing cost growth without sacrificing the health gains made possible by improved health care.

QUESTIONS AND PROBLEMS

1. **E** Navigate to the Excel sheet for Table A3, “Percentage of People Without Health Insurance Coverage by State: 2010, 2018, and 2019” of the Census Bureau report “Health Insurance Coverage in the United States: 2019,” found at
<https://www.census.gov/library/publications/2020/demo/p60-271.html>.
 - a. Using this spreadsheet, calculate the mean uninsured rate for Medicaid Expansion states, and states that have not expanded, for both 2010 and 2019. (Include the District of Columbia in your calculations.)
 - b. What does the result in part (a) indicate about take-up versus crowd-out during the recent Medicaid expansions enabled by the Affordable Care Act?
2. **E** In 1997, the Children’s Health Insurance Program was introduced, and through expansions over the next 15 years this program substantially increased children’s access to public health insurance beyond the traditional limits of the Medicaid program. How do you think this affected the job mobility of low- to middle-income parents? How could you empirically investigate this question?
3. **E** The Oregon Medicaid Experiment, described in [Finkelstein et al \(2012\)](#), found that individuals receiving Medicaid had improved mental health and were also under reduced financial strain, having significantly reduced medical debt and fewer bills in collections. Explain how these results demonstrate that some of the primary welfare gains from Medicaid are due to its role as traditional social insurance that provides consumption smoothing.
4. Explain why take-up rates—the fraction of eligible individuals who enroll in the program—are so much higher for Medicare than for Medicaid.
5. Describe the main components of Massachusetts’ health care reform. In what ways did Massachusetts provide a model for the Affordable Care Act? How was Massachusetts’ experience with reform similar and different to national experience with reform under the Affordable Care Act?
6. Congressman Brown supports pre-existing conditions protections introduced by the Affordable Care Act. These protections force insurers to cover health problems arising from pre-existing conditions and forbid them from charging people with such health conditions higher prices for insurance. However, he opposes the individual mandate and subsidies introduced by the ACA, arguing that the individual mandate inhibits individual freedom, and subsidies are too costly for the government. He ~~supported the 2017 repeal of the individual mandate, and currently~~

- supported the 2017 repeal of the individual mandate, and currently supports going even further, wishing to repeal the ACA in its entirety. As such, he has introduced legislation to fully repeal the ACA and has introduced companion legislation containing pre-existing condition protections alone. What do you think might happen if he was successful?
7. What are the similarities between Medicare vouchers and education vouchers (described in [Chapter 11](#))? What are the differences?
 8. Since the 1960s, the percentage of U.S. GDP spent on National Health Expenditure has risen from around 5% to around 18% (as of 2016). Is this rapidly expanding share of total GDP of the health sector necessarily evidence of wasteful health care spending? Why or why not?
 9. One important goal of the Affordable Care Act was to address rising health care spending. Consult the free-access article “National Health Care Spending In 2016: Spending and Enrollment Growth Slow after Initial Coverage Expansions,” published in *Health Affairs*, at <https://www.healthaffairs.org/doi/10.1377/hlthaff.2017.1299>. This article reports aggregate statistics on national health expenditure using data from the Centers for Medicare and Medicaid Services.
 - a. Consult Exhibit 1. What was the annual growth in national health expenditure in the years prior to and after the Affordable Care Act? In 2014–2015, immediately after the ACA’s main provisions came into effect, what were the areas with greatest growth in expenditure?
 - b. Consult Exhibit 2. For each category of insurance coverage in the exhibit (private insurance, Medicare, and Medicaid), what has happened in the years after the ACA to enrollment growth and per-enrollee expenditure growth?
 - c. Discuss how these tables shed light on the impact the ACA has had on overall growth in national health expenditure. How do your findings relate to the various cost-control measures introduced by the ACA that were discussed in the chapter?
 10. What explains the change in long-term care hospital discharge patterns depicted in [Figure 16-2](#)? What are the implications for policy?
 11. The fact that such a large percentage of U.S. health care costs is spent on people in their last six months of life has led some to claim that the American health care system is “wasteful.” Why might this be an overgeneralization?
 12. In 2017, partial repeal of the Affordable Care Act was enacted by Congress as a part of a larger tax reform package that made many substantial changes to the tax code. Why does the manner in which it was passed

complicate efforts to evaluate the effect of partial ACA repeal on health care costs, coverage, and the overall economy?

ADVANCED QUESTIONS

13. Describe the empirical evidence for a relationship between expanding Medicaid eligibility and improving health outcomes. As depicted in [Figure 16-1](#), how strong is the empirical evidence for each of Steps 1–3 in this relationship?
14. Suppose all individuals have a utility function of $U = \sqrt{C}$, where C is the amount of consumption a person has in any given period. There are two types of people: those at high risk, with a 4% chance of accident, and those at low risk, with a 2% chance of accident. Half of individuals are high risk, and half are low risk. The income of all individuals is \$40,000 per year, and a catastrophic accident costs \$30,000 in the year of the accident.
 - a. Calculate the actuarially fair insurance premium for each type of person, assuming the insurance company can identify who is high risk and who is low risk.
 - b. How much will a high-risk individual be willing to pay for insurance? A low-risk individual?
 - c. Suppose a law passes requiring community rating, and all insurance contracts must be priced the same. Insurers are forward-looking, and know about the composition and preferences of people in the insurance market, but are forbidden from using that information to set prices. What would be the price for insurance, and who would purchase insurance?
 - d. The legislature is considering two possible additional regulations on top of community rating: a mandate, and subsidies. The mandate would charge anyone failing to purchase insurance \$100. The subsidies would give anyone purchasing insurance \$100 to offset the price of the insurance. Under this scheme, what would the price be for insurance and who would purchase it?
 - e. How do the results above relate to Massachusetts' and the ACA health care reform strategies?
15.  In response to the State Children's Health Insurance Program in 1997, 37 states (including the District of Columbia) expanded Medicaid coverage

to children in families below 200% of the poverty line, and even higher thresholds in some states.

- a. In some of these states, the eligibility expansions have covered all children. How would you design a quasi-experimental analysis to evaluate the impact of these expansions?
 - b. In other states, the eligibility expansions only covered certain age groups of children. How could you design a quasi-experimental analysis to evaluate the impact of these expansions? How could you make this more convincing than the evaluation in part (a)? Explain.
16. After the Medicare program adopted the Prospective Payment System (PPS), researchers observed that people tended to receive less care for any given diagnosed condition.
- a. One explanation for this finding is that the PPS provides incentives to provide lower levels of treatment for any given diagnosis. Why would PPS provide these incentives?
 - b. Another explanation for this finding is that PPS offers incentives for physicians to diagnose marginal health conditions as more serious than they are. Why would PPS provide these incentives?
 - c. Since this reduction in quantity of care was not accompanied by a reduction in observable health outcomes, what, if anything, can you infer about the efficiency of the Medicare program *before* the policy change? Explain your answer.
17. One feature of Medicare coverage is that individuals are responsible for 20% of their Part B (primary physician) costs, without limit. Individuals have traditionally purchased Medi-gap policies that cover this gap by paying for the out-of-pocket costs not covered by Medicare. But some Medi-gap policies did not cover this 20% copayment.
- [Finkelstein \(2002\)](#) studied the effects of a federal mandate that Medi-gap plans cover this 20% copayment. She found that this mandate would lead fewer individuals to buy Medi-gap coverage.
- a. Why would the mandate lower the demand for Medi-gap coverage?
 - b. What do you think would be the net effect of this policy on the costs of the Medicare program itself?
18. In 1981, the federal government passed a law that gave states permission to change the structure of their Medicaid program. States could now, if they wished, require Medicaid beneficiaries to enroll in a Medicaid “managed care organization” (MCO), as long as Medicaid recipients were offered a choice of several plans. Medicaid recipients would be required

offered a choice of several plans. Medicaid recipients would be required to receive their medical care only through their MCOs. The MCOs would receive fixed, regular payments from the state and, in return, would cover the medical expenses of their Medicaid enrollees.

- a. Using what you know about Medicaid and managed care, explain several reasons why policy makers might support the requirement that Medicaid beneficiaries enroll in MCOs.
 - b. Again, applying what you know about Medicaid and managed care, how do you think this requirement would affect the decision of people who are eligible to enroll in Medicaid? Be specific about which individuals eligible for Medicaid are likely to change or not change their decision to enroll.
 - c. How might this requirement affect overall access to care for Medicaid-eligible patients?
19. The current government-provided system in the country of Puceland provides free health insurance for all children but to no adults. There are two types of adults in Puceland: high earners and low earners. All the 100,000 high earners receive insurance coverage through their employer, but only half of the 100,000 low earners do. The remaining adults are uninsured.

You are hired to analyze the effectiveness of a proposed plan to offer coverage to all low earners. You have read the economics literature in Puceland and your best estimates are as follows: (1) only 80% of uninsured workers who are offered government health insurance will choose to enroll; (2) 60% of currently insured low earners work at firms that will drop insurance coverage for them after the policy change; the other 40% will remain in their current employer-provided plan; (3) 10% of high earners will choose to become low earners (at firms who do not offer health insurance) and take up the government insurance once they can get it.

- a. Estimate the increase in the number of insured adults.
- b. Estimate the dollar cost per additionally insured adult. Why is it so much higher than \$5,000?
- c. Suppose that, without access to any insurance, each adult has a 5% chance of dying in a given year. Access to government-provided health insurance reduces the chance to 3%, while access to employer-provided health insurance reduces it to 2%. If it costs the government \$5,000 to provide health insurance per year to an adult, estimate the dollar cost

of the program per life saved.

The  icon indicates a question that requires students to apply the empirical economics principles discussed in [Chapter 3](#) and the Empirical Evidence boxes.
arbuz/Shutterstock

DO NOT COPY
reibaker@live.unc.edu