Research Report

The RAND Health Care Payment and Delivery Simulation Model (PADSIM)

Concepts, Methods, and Examples

Chapin White, Jodi L. Liu, Mikhail Zaydman, Sarah A. Nowak

Peter S. Hussey

RAND CORPORATION

For more information on this publication, visit www.rand.org/t/rr1428

Published by the RAND Corporation, Santa Monica, Calif.

Preface

This report describes concepts and mechanics of RAND's Health Care Payment and Delivery Simulation Model (PADSIM). The purpose of this report is to provide analysts, both within and outside RAND, background on the motivation for building the model; give them an understanding of the conceptual underpinnings of the model; and provide an overview of how to operate the model. We anticipate that this report will be updated as the model continues to be revised and applied to new research questions.

Funding for this study was provided by philanthropic contributions from RAND supporters and income from operations.

Contents

Figures and Tables

Figures

Tables

Acknowledgments

The authors are extremely grateful to Susan L. Marquis and C. Richard Neu for their support, and to Jeffrey Wasserman for his enthusiasm and guidance. Several RAND staff and affiliates contributed to the development of PADSIM, including Christine Eibner, Carolyn Rutter, and Paul Ginsburg. We benefited greatly from the thoughtful comments provided by David Auerbach and Srikanth Kadiyala on an earlier version of this report. Finally, we would like to thank our RAND colleagues Steve Oshiro and James Torr for their assistance with this manuscript.

1. Introduction

In late 2014, the RAND Corporation began to develop the Health Care Payment and Delivery Simulation Model (PADSIM) to build RAND's capacity to analyze the impacts of current and future reforms to provider payment policy. These reforms include changes in how much providers are paid for their services and changes in the degree to which payments are determined by the volume of services. Payment reforms are clearly a high priority for public and private purchasers, and we expect payment reform to remain in the spotlight for the foreseeable future. PADSIM offers a systematic framework for quantifying such reforms and providers' responses to them.

Why Do We Need Another Model?

Countless models have been developed to analyze the impacts of health care reform proposals.[1] These models, and their outputs, are extremely important to the policymaking process: They provide a preview of the impacts of proposed policies, they quantify trade-offs, and, in some cases, they feed directly into federal legislative and regulatory processes. Prominent examples include

- the RAND COMPARE model[2]
- the Congressional Budget Office's (CBO's) Health Insurance Simulation Model (HISIM)[3]
- the Urban Institute's Health Insurance Reform Simulation Model (HIRSM)[4]
- the Lewin Group's Health Benefits Simulation Model (HBSM)[5]
- the Centers for Medicare and Medicaid Services' (CMS's) Office of the Actuary Health Reform Model (OHRM).[6]

---

[1] For an excellent overview of modeling approaches, see Sherry Glied and Nicholas Tilipman, "Simulation Modeling of Health Care Policy," *Annual Review of Public Health*, Vol. 31, No. 1, 2010, pp. 439–455.

[2] Federico Girosi, Amado Cordova, Christine Eibner, Carole Roan Gresenz, Emmett Keeler, Jeanne Ringel, Jeffrey Sullivan, John Bertko, Melinda Beeuwkes Buntin, and Raffaele Vardavas, *Overview of the COMPARE Microsimulation Model*, Santa Monica, Calif.: RAND Corporation, WR-650, 2009; Christine Eibner and Carter C. Price, *The Effect of the Affordable Care Act on Enrollment and Premiums, With and Without the Individual Mandate*, Santa Monica, Calif.: RAND Corporation, TR-1221-CMF, 2012.

[3] Congressional Budget Office, *CBO's Health Insurance Simulation Model: A Technical Description*, Washington, D.C., October 2007; Congressional Budget Office, *Key Issues in Analyzing Major Health Insurance Proposals*, Washington, D.C., December 2008; Congressional Budget Office, *Cost Estimate for the Amendment in the Nature of a Substitute for H.R. 4872, Incorporating a Proposed Manager's Amendment Made Public on March 20, 2010*, Washington, D.C., March 20, 2010.

[4] Linda J. Blumberg, Yu-Chu Shen, Len M. Nichols, Matthew Buettgens, Lisa C. Dubay, and Stacey McMorrow, *The Health Insurance Reform Simulation Model (HIRSM): Methodological Detail and Prototypical Simulation Results*, Washington, D.C.: Urban Institute, July, 2003.

[5] The Lewin Group, *The Health Benefits Simulation Model (HBSM): Methodology and Assumptions*, March 31, 2009.

[6] Centers for Medicare and Medicaid Services, *Projections of National Health Expenditures: Methodology and Model Specification*, Washington, D.C., September 3, 2014.

In general, these models are designed to focus on health insurance coverage and premiums, including the impacts of proposals to expand eligibility for public coverage, provide subsidies for insurance, or penalize those who go uninsured. The design of these models reflects the long-standing interest of U.S. policymakers in expanding coverage and reducing the number of uninsured.

PADSIM is designed differently from these existing models—it focuses on health care providers and their responses to changes in "payment policy," meaning the level of payments providers receive, and the relationship between the payments they receive and the services they provide. The development of PADSIM, and its focus on provider payment policy, reflects four recent developments:

1. *Coverage has been expanded*. The existing health reform models—COMPARE, HISIM, HIRSM, and HBSM—were tailored to inform the design of the coverage expansions in the Patient Protection and Affordable Care Act (ACA) and predecessor proposals from the 1990s and 2000s. The key question in those policy debates was how to effectively expand health insurance coverage while limiting government outlays to a reasonable level. That question has been answered, for now at least, by the enactment and implementation of the ACA. We recognize that the ACA has been seriously challenged in the courts, and there have been countless proposals, and even legislation passed in the House of Representatives, to repeal the law. Notwithstanding, we view the ACA's general approach—Medicaid for the poor, tax credits for nongroup coverage for those with low or moderate incomes, and an individual mandate to glue the market together—as a durable one. Important policy questions will continue to emerge as the ACA is implemented and replacements are proposed, but the existing modeling capacity is well suited to addressing these implementation questions. In contrast, reducing the growth in health care spending while improving the value of care provided is a central problem for policymakers and health care decisionmakers, and there is little modeling capacity and limited empirical information for various policy options.
2. *Payment reform is a work in progress.* Policymakers have increasingly emphasized reforms to payment to providers and the role of those reforms in improving the efficiency of the U.S. health care system. Emblematic of this shift is the announcement by Secretary Sylvia Mathews Burwell of the U.S. Department of Health and Human Services (DHHS) of a goal to shift the vast majority of Medicare's payments to some form of "value-based" payment over the next few years.[7] Unlike the approach to coverage expansions in the ACA, we view provider payment reform as a highly unsettled policy area, with a huge variety of proposals in play and reforms in the early stages of implementation.
3. *The evidence base has grown*. In the 1990s, the knowledge base regarding provider responses to changes in payment policy was relatively thin and focused mainly on hospital responses to prospective payment,[8] physician practice patterns in a managed care environment,[9] and the

---

[7] Sylvia M. Burwell, "Setting Value-Based Payment Goals—HHS Efforts to Improve U.S. Health Care," *New England Journal of Medicine*, 2015, pp. 897–899.

[8] Judith Feder, Jack Hadley, and Stephen Zuckerman, "How Did Medicare's Prospective Payment System Affect Hospitals?" *New England Journal of Medicine*, Vol. 317, No. 14, October 1, 1987, pp. 867–873; David M. Cutler, "The Incidence of Adverse Medical Outcomes Under Prospective Payment," *Econometrica,* Vol. 63, No. 1, January 1995, pp. 29–50.

[9] Arnold M. Epstein, Colin B. Begg, and Barbara J. McNeil, "The Use of Ambulatory Testing in Prepaid and Fee-for-Service Group Practices," *New England Journal of Medicine*, Vol. 314, April 24, 1986, pp. 1089–1094; G. M. Clancy

so-called physician "volume offset."[10] In more recent years, a number of studies have been published on provider responses to changes in payment policy, including

- *production costs*, i.e., how providers adjust their costs of production in response to changes in payment policy[11]
- *supply curve*, i.e., estimates of changes in volume of services in response to a change in the payment rate for those services[12]
- *spillovers*, i.e., estimates of the effects of changes in payments from one payer for one type of provider or patient affect payments from other payers or services provided to other patients or in other settings.[13]

The profusion of recent evidence provides a much stronger basis for behavioral estimates in a provider-focused simulation model.

4. *The recent slowdown in health care spending growth*. A growing mystery is why health spending growth has slowed markedly in recent years, significantly improving government budget forecasts. The Great Recession can account for some, but not all, of the slowdown, and analysts generally agree that other factors seem to be in play. One likely suspect is the ACA, which has stimulated widespread changes in payment policy and innovation across the health care delivery system. Although the public's attention has focused on the ACA's

---

and B. E. Hillner, "Physicians as Gatekeepers: The Impact of Financial Incentives," *Archives of Internal Medicine*, Vol. 149, 1989, pp. 917–920; the "Results at the Health Maintenance Organization" section in Joseph P. Newhouse and the Insurance Experiment Group, *Free for All? Lessons from the RAND Health Insurance Experiment*, Cambridge, Mass.: Harvard University Press, 1993.

[10] Nguyen Xuan Nguyen and Frederick William Derrick, "Physician Behavioral Response to a Medicare Price Reduction," *Health Services Research*, Vol. 32, No. 3, August 1997, pp. 283–298; Health Care Financing Administration, Office of the Actuary, *Physician Volume & Intensity Response*, Washington, D.C., August 13, 1998.

[11] Chapin White and Vivian Yaling Wu, "How Do Hospitals Cope with Sustained Slow Growth in Medicare Prices?" *Health Services Research*, Vol. 49, No. 1, February 2014; Chapin White, "Medicare's New Prospective Payment System for Skilled Nursing Facilities: Effects on Staffing and Quality of Care," *Inquiry*, Vol. 43, No. 1, Winter 2005/2006, pp. 351–366.

[12] Chapin White and Tracy Yee, "When Medicare Cuts Hospital Prices, Seniors Use Less Inpatient Care," *Health Affairs*, Vol. 32, No. 10, October 2013, pp. 1789–1795; Daifeng He and Jennifer M. Mellor, "Hospital Volume Responses to Medicare's Outpatient Prospective Payment System: Evidence from Florida," *Journal of Health Economics*, Vol. 31, 2012, pp. 730– 743; Jack Hadley and James D. Reschovsky, "Medicare Fees and Physicians' Medicare Service Volume: Beneficiaries Treated and Services per Beneficiary," *International Journal of Health Care Finance and Economics*, Vol. 6, No. 2, June 2006, pp. 131–150; Congressional Budget Office, *Factors Underlying the Growth in Medicare's Spending for Physicians' Services*, Washington, D.C., June, 2007; Jeffrey Clemens and Joshua D. Gottlieb, "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" *American Economic Review*, Vol. 104, No. 4, 2014, pp. 1320–1349; Abe Dunn and Adam Hale Shapiro, *Physician Market Power and Medical-Care Expenditures*, Washington, D.C.: U.S. Bureau of Economic Analysis, April 26, 2012; Chapin White and Nguyen Nguyen, "How Does the Volume of Post-Acute Care Respond to Changes in the Payment Rate?" *Medicare & Medicaid Research Review*, Vol. 3, No. 1, 2011, pp. E1–E22.

[13] Daifeng He and Jennifer M. Mellor, "Do Changes in Hospital Outpatient Payments Affect the Setting of Care?" *Health Services Research*, Vol. 48. No. 5, October 2013; Chapin White "Cutting Medicare Hospital Prices Leads to a Spillover Reduction in Hospital Discharges for the Nonelderly," *Health Services Research*, Vol. 149, No. 5, 2014, pp. 1578–1595; Chapin White, "Contrary to Cost-Shift Theory, Lower Medicare Hospital Payment Rates for Inpatient Care Lead to Lower Private Payment Rates," *Health Affairs,* Vol. 32, No. 5, May 2013, pp. 935–943.

expansion of insurance coverage, the supply-side changes in the ACA may turn out to have an even more significant impact.

Overview of PADSIM's Inputs and Outputs

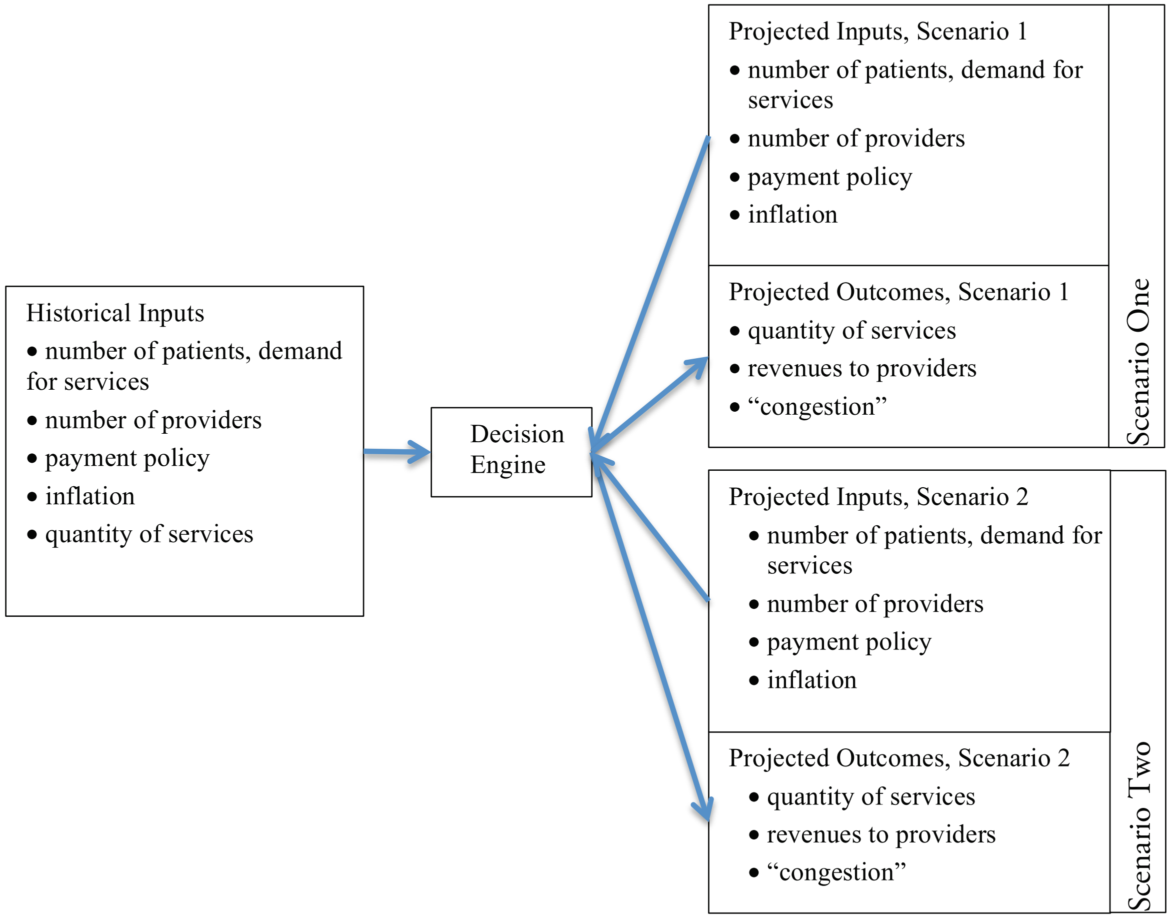
Figure 1.1 provides a very high-level, simplified view of PADSIM's inputs and outputs. PADSIM's key outputs are projected quantities of health care services provided, the revenues paid to providers for those services, and a level of "congestion," which is a measure of the degree to which patients' demand for services exceeds providers' desired output. To generate those outputs, PADSIM uses two types of inputs. The first is historical data on the number of patients and their demand for health care services, the number of providers, provider payment policy, and the actual quantity of services provided. The second is projections of the number of patients and their demand for health care services, and payment policy.

PADSIM uses what we refer to as a *decision engine*, which is software code that embodies a set of assumptions regarding the behavior of health care providers and patients. The decision engine is first applied to the historical data, and the quantity predicted by the decision engine is compared with the actual historical quantity of services—the difference is the *historical residual*. The decision engine is then applied to the inputs for a given projection scenario, and the decision engine produces a predicted level of output of health care services that incorporates patient demand, provider payment policy, and the historical residuals.

PADSIM becomes useful when we define multiple projection scenarios and compare projected outcomes under those different scenarios. For example, we can define one scenario based on projections of payment policy under the Sustainable Growth Rate formula (the "SGR" scenario), and another based on projections of payment policy under the Medicare Access and CHIP Reauthorization Act of 2015 (the "MACRA" scenario).[14] (We explain these scenarios in more detail in Chapter Five.) The difference between projected outcomes under those two scenarios represents the simulated impact of repealing the SGR and replacing it with MACRA.

[14] Public Law 114-10, Medicare Access and CHIP Reauthorization Act of 2015, April 16, 2015.

Figure 1.1. PADSIM's Inputs and Outputs



SOURCE: Authors' analysis.

2. PADSIM's Structure and Key Concepts

Unit of Analysis and Scope of the Model

PADSIM is, at its heart, a model of the supply and demand for health care services. The "unit of analysis" refers to the granularity of the model, or, more specifically, the level at which supply and demand are quantified and assumed to be in equilibrium. In the current (2014–2015) version of PADSIM, the unit of analysis is the combination of state, year, provider type (physician versus hospital), and coverage type (Medicare, Medicaid, private group, nongroup, and uninsured). Future versions of PADSIM will likely use more-granular units of analysis, such as counties rather than states, and differentiate between primary care versus specialist physician services.

To illustrate, the current version of PADSIM quantifies the demand for, and supply of, physician services among Medicare beneficiaries in Virginia in 2016. The current version of PADSIM does *not* quantify the demand for, or supply of, physician services within specific counties within Virginia, or among demographic subgroups, or in specific months. PADSIM also does *not* quantify demand for, or supply of, specific types of physician services (e.g., imaging versus evaluation and management) or services provided by individual physicians or types of physicians (e.g., primary care physicians versus specialists). We refer to this approach as a *semiaggregated* model, because the unit of analysis represents aggregated groups of patients and providers, but not aggregated all the way up to the national level.

The scope of the model refers to the range of units of analyses that are included in the model. The scope and unit of analysis in PADSIM are specified in Table 2.1.

Table 2.1. Scope and Unit of Analysis

Unit of Analysis	Scope	Notes
State	50 U.S. states plus the District of Columbia	Excludes U.S. territories.
Year	1997–2026	
Provider type	Offices of physicians	Corresponds to NAICS 6211, includes physician services plus ancillary services, such as imaging and laboratory services, provided in physician offices.
	General medical and surgical hospitals	Corresponds to NAICS 6221, includes hospital inpatient acute care services and outpatient services, and excludes hospital-based post-acute care.
Health insurance coverage	Medicare	Includes all individuals whose primary source of coverage is Medicare, including individuals dually eligible for Medicare and Medicaid.
	Medicaid	Includes enrollees in the Children's Health Insurance Program (CHIP) and other state and local programs for the indigent.
	Private group	Includes enrollees in health plans for government employees, including members of the military and their dependents, retiree health plans, and other sources of coverage not included elsewhere.
	Nongroup	Includes enrollees in Marketplace plans, and other individuals whose primary source of coverage is a nongroup plan. Excludes individuals who purchase nongroup supplemental coverage.
	Uninsured	Includes individuals who are not enrolled in a source of health insurance coverage, regardless of whether they are eligible or not.

NOTE: NAICS = North American Industry Classification System.

In principle, PADSIM could be designed to allow for a huge range of possible interactions among the various units of analysis. For example, physician payment policy in one state could result in physician migration and thereby affect the number of physicians practicing in other states. To keep the model tractable, PADSIM is designed to allow only two possible interactions among the units of analysis:

- *Spillovers among coverage types*: The demand for services and the payment policy for patients with one type of coverage can have spillover impacts on utilization of services among other patients within a state-year-provider type; and
- *Spillovers among provider types*: The payment policy for one provider type (e.g., physicians) can affect utilization of services of a different provider type (e.g., hospitals) within a state-year. This type of spillover can occur if, for example, physicians are financially rewarded based on the utilization of hospital services among their panel of patients.

In the current version of PADSIM, these two types of spillovers occur only within a state-year, which greatly simplifies the operation of the model.

Crossing of State Boundaries

The vast majority of physician and hospital services are provided within-state, meaning that the state where the service is provided is the patient's state of residence and also the state of the provider's main practice location. In some cases, however, patients travel to neighboring states to receive medical care, or they receive services while on vacation. For example, large teaching hospitals in

Massachusetts serve many residents of neighboring states in New England. Physicians may also travel to provide services in multiple locations, some of which may cross state lines. These border crossings are particularly common in the New York City metropolitan area and the Washington, D.C., metropolitan area.

In PADSIM, patient demand for health care services is simulated based on the number and characteristics of the residents of a state, while provider supply is simulated based on the number of providers in a state and the payment policy those providers face. Border crossing is, implicitly, taken into account in the use of residuals to autocalibrate the model (described later in this chapter). For example, the quantity of hospital services provided by hospitals in Massachusetts is larger than would be expected based solely on the demand for hospital services among residents of Massachusetts, and that difference, or residual, is included in projections of hospital services provided in Massachusetts. But border crossing is not taken into account in measuring payment policy as it affects providers in a state or in measuring the demand for services.

Defining the Volume of Services

Conceptually, the quantity of services in PADSIM equals the number of services adjusted for the intensity of those services.

- *Hospital quantity*. Hospitals produce both inpatient and outpatient services. The quantity and intensity of inpatient services are relatively straightforward to measure—the number of inpatient stays can be used as a measure of raw quantity and the average diagnosis-related group (DRG) weight can be used as a measure of intensity. (Inpatient stays are classified into DRGs for payment purposes based on similarities in clinical diagnoses and procedures, and similarity in resource requirements.) Measuring the quantity and intensity of hospital outpatient visits is much more challenging. This is partly because outpatient visits are highly variable in their intensity, ranging from simple blood draws to outpatient surgeries. The methodologies for quantifying intensity in the hospital outpatient setting are far less developed than in the inpatient setting. PADSIM does not currently differentiate among different types of services provided within the hospital setting, and, therefore, we defined a single metric of hospital service quantity. We measured the quantity of hospital services in units of intensity-weighted stay-equivalents, calculated as follows:
  - the number of inpatient hospital stays multiplied by the mean Medicare DRG weight for those stays, *plus*
  - outpatient discharge-equivalents, which equals the quantity of outpatient services expressed in inpatient stay equivalents. We first calculated the ratio of costs that hospitals incur in providing outpatient services to the costs that hospitals incur in providing inpatient services using Medicare Cost Reports. This ratio provides a measure of the relative output of outpatient and inpatient services provided by each hospital. This ratio was then multiplied by the quantity of DRG-weighted inpatient stays, which produces a measure of outpatient quantity in units of inpatient stay equivalents. These methods

follow the spirit of American Hospital Association's concept of "adjusted days" and have been used in a previous study.[15]

- *Physician quantity*. In the physician office setting, the quantity of services equals the sum of the relative value units (RVUs) for all services provided. RVUs are designed to reflect the intensity of resources needed to produce various types of physician services, including the physician's work plus the expense of maintaining the office and purchasing supplies.

Payment Policy

PADSIM's key innovation is to systematically quantify provider payment policy and incorporate providers' behavioral responses into the equilibrium concept. This requires a high degree of simplification while, it is hoped, retaining the key phenomena of interest. In the real world, provider payment policies—the arrangements that govern payments to providers—vary in ways that are almost unimaginably complex. In the traditional Medicare program, the rules that govern payments to providers consist of base payment rates, case-mix adjustments, adjustments for local market conditions, adjustments for different provider types, outlier payments, exceptions for special types of providers or services, pay-for-performance bonuses, special rules for participants in voluntary payment demonstrations, teaching adjustments, and on and on.[16] In principle, all of these payment arrangements in traditional Medicare are described in detail in federal legislation and regulations, although, in practice, many of the details will be comprehensible only to a narrow set of experts. Medicare also operates a major program—Medicare Advantage—in which beneficiaries enroll in private health plans. The payment arrangements in Medicare Advantage are opaque, and researchers studying payment arrangements have had to rely on laborious qualitative research to establish even the most basic facts.[17] Payment arrangements in state Medicaid programs exhibit a level of complexity that is similar to Medicare, but with the added complexity of varying profoundly from one state to another. In private health plans, two additional challenges arise in attempting to understand payment arrangements: Plans enter into individualized contracts that vary from provider to provider, and those contracts are treated as trade secrets.

The approach taken in PADSIM is to summarize the payment arrangements between a health plan and a provider type using two concepts:

- *Payment rate*, meaning the average revenue to the provider per volume of services provided. The concept of the payment rate is relatively straightforward—for example, in California in 2013, Medicare payments to physicians were $39.95 per RVU.

---

[15] For a description of the calculation of discharge-equivalents, see Chapin White and Vivian Yaling Wu, "How Do Hospitals Cope with Sustained Slow Growth in Medicare Prices?" *Health Services Research,* Vol. 49, No. 1, February 2014, pp. 11–31.

[16] For a concise overview of one of Medicare's traditional payment systems, with some indication of the level of complexity, see Medicare Payment Advisory Commission, *Hospital Acute Inpatient Services Payment System*, Washington, D.C., October 2014.

[17] R. A. Berenson, J. H. Sunshine, D. Helms, and E. Lawton, "Why Medicare Advantage Plans Pay Hospitals Traditional Medicare Prices," *Health Affairs,* Vol. 34, No. 8, 2015, pp. 1289–1295.

- *Prospectiveness*, meaning the degree to which payments to providers' revenues are delinked from the costs providers incur. The concept of prospectiveness is less straightforward and requires more explanation.

Prospectiveness, in PADSIM, is an index that is defined using two anchors:

- *Cost reimbursement—prospectiveness equals 0*. In a pure cost reimbursement arrangement, providers are reimbursed by the health plan for whatever costs they incur in treating the plan's patients. In this case, the elasticity of provider revenues with respect to costs incurred equals 1.
- *Global capitation—prospectiveness equals 1*. In a pure global capitation arrangement, providers are paid a fixed amount for each patient assigned to them, regardless of the quantity or intensity of the services provided to those patients. In this case, the elasticity of provider revenues with respect to costs incurred equals 0.

It is important to point out that these anchor points represent extreme scenarios that are rare or nonexistent in the real world. Real-life payment arrangements almost invariably fall somewhere on the continuum between cost reimbursement and global capitation.
The general definition of prospectiveness is

$$Pro_{s,c,y} = 1 - \frac{\partial R_{s,c,y} / \partial C_{s,c,y}}{R_{s,c,y} / C_{s,c,y}}$$

where

- $R_{s,c,y}$ is the revenues to physicians and hospitals in state *s* for treating patients with coverage *c* in year *y*—revenues include payments from the health plan and out-of-pocket payments from the patient, and include any performance bonuses or withholds
- $C_{s,c,y}$ is the costs incurred by physicians and hospitals in state *s* for treating patients with coverage *c* in year *y*—costs include wages paid to staff, equipment and supplies, earnings to owners
- $\partial R_{s,c,y} / \partial C_{s,c,y}$ is the marginal revenue to physicians and hospitals with respect to costs incurred.

To be operationalized in PADSIM, the general definition of prospectiveness requires two key refinements:

- *Own- versus cross-provider prospectiveness*. So far, the discussion of prospectiveness has not differentiated between different types of providers, which glosses over two important phenomena. First, a health plan may capitate payments to one type of provider while paying another type of provider separately for each service—a typical example is a health plan paying a capitated rate to a physician practice, while paying hospitals separately for each service they provide. In that case, the "own-provider" prospectiveness will differ for physicians versus hospitals. Second, in some payment arrangements, the revenues received by one type of provider will depend in part on the costs of services provided by a different type

of provider—those arrangements affect what we refer to as *cross-provider prospectiveness*. The rationale for including cross-provider prospectiveness in the model is that one type of provider may change care delivery in a way that reduces the costs of care provided by other providers. For example, Medicare's value-based payment modifier (VBPM) for physicians applies a bonus or penalty to physician payment rates based, in part, on the cost of services provided by hospitals and other types of health care providers.[18] In response, physicians may take steps to reduce the frequency of hospitalization of their patients. A second example is bundled payment for an episode of hospital and post-acute care, whereby hospitals are paid in part based on the total cost of the episode. In response, hospitals may seek to minimize the costs of post-acute care through selection of low-cost post-acute care providers or reduced use of institutional post-acute care. A third, and more extreme, example is a global capitation arrangement in which a physician organization receives a global payment from a health plan for a panel of patients, and then the physician organization in turn pays hospitals and other providers for services they provide to those patients.

To recognize these interrelationships among providers, we generalize the concept of prospectiveness as follows:

$$Pro_{s,c,p1,p2,y} = 1 - \frac{\partial R_{s,c,p1,y} / \partial C_{s,c,p2,y}}{R_{s,c,p1,y} / C_{s,c,p2,y}}$$

where $p1$ is provider type 1, and $p2$ is provider type 2.

This allows us to differentiate between (1) *system-wide prospectiveness*, which treats all providers jointly, (2) *own-provider prospectiveness*, which relates to how revenues to one type of provider vary depending on the services they provide (i.e., $p1$ and $p2$ are the same provider type), and (3) *cross-provider prospectiveness*, which relates to how revenues to one type of provider vary depending on the costs of services provided by a different type of provider (i.e., $p1$ and $p2$ are different provider types).

Revenues, in this context, refer to the revenues to one type of provider, net of any payments to other types of providers.

- *Expectations.* Health care providers face uncertainties in the types of patients they will treat and the services they will provide. As a result, all of the financial variables used to define prospectiveness—revenues, costs, and the elasticity of revenues with respect to costs—may be defined in two different ways: (1) based on the provider's expectation (i.e., *ex ante*, before services are provided), or (2) based on the realized outcome (i.e., *ex post*, after services are provided). For example, Medicare's inpatient hospital payment system includes an outlier provision for unusually high-cost cases that reimburses hospitals for 80 percent of costs above a fixed-loss threshold. The outlier provision, *ex ante*, increases the elasticity of revenues with respect to costs for all inpatient admissions, because any admission could, potentially, generate outlier payments. But, *ex post*, only a small share of admissions actually do generate outlier payments. Similarly, *ex ante*, all accountable care organizations (ACOs) in Medicare

[18] Centers for Medicare and Medicaid Services, *2014 Measure Information About the Per Capita Costs for All Attributed Beneficiaries Measure, Calculated for the Value-Based Payment Modifier Program*, Washington, D.C., April 2015.

are potentially eligible for bonus payments if they meet quality and spending targets, although, *ex post*, only some ACOs will actually receive bonuses.

The approach taken in PADSIM is to define prospectiveness for a state-year-coverage type based on providers' expectations over the course of a year regarding their revenues, costs, and marginal revenues with respect to costs for patients with that coverage type. We assume that providers' expectations match actual historical outcomes at the aggregate level. For example, in the Medicare Shared Savings Program (MSSP) in 2014, 92 out of 333 ACOs qualified for shared savings. In calculating the boost in prospectiveness due to MSSP, we would then assume that participating providers anticipate roughly a one-in-four probability (i.e., 92 out of 333) that they will qualify for shared savings and roughly a three-in-four probability (i.e., 241 out of 333) that they will not.

Table 2.2 illustrates the calculation of own- and cross-provider prospectiveness under four hypothetical payment policies. For this illustration, we considered only revenues to physicians and hospitals, and we assumed that physician revenues accounted for one-quarter of total revenues to physicians and hospitals.

Table 2.2. Illustration of Calculation of Own- and Cross-Provider Prospectiveness Under Four Example Payment Policies

					Payment Policy			
				Notation	Cost Reimbursement	Physicians Receive Capitation, Hospitals Receive Fee-for-Service	Physicians Receive Capitation, Hospitals Receive Separate Capitation	Physicians Receive Global Capitation, Physicians Pay for Hospital Care Out of Global Amount
Marginal revenues to …	all providers	… with respect to costs incurred by …	all providers	$\partial R_{s,c,y} / \partial C_{s,c,y}$	1	0.75	0	0
	physicians		physicians	$\partial R_{s,c,phys,y} / \partial C_{s,c,phys,y}$	1	0	0	0
	physicians		hospitals	$\partial R_{s,c,phys,y} / \partial C_{s,c,hosp,y}$	0	0	0	-1
	hospitals		hospitals	$\partial R_{s,c,hosp,y} / \partial C_{s,c,hosp,y}$	1	1	0	1
	hospitals		physicians	$\partial R_{s,c,hosp,y} / \partial C_{s,c,phys,y}$	0	0	0	0
Prospectiveness	system-wide (all providers)			$Pro_{s,c,y}$	0	0.25	1	1
	physician, own-prospectiveness			$Pro^{own}_{s,c,phys,y}$	0	1	1	1
	physician-hospital cross-prospectiveness			$Pro^{cross}_{s,c,hosp,phys,y}$	1	1	1	4
	hospital, own			$Pro^{own}_{s,c,hosp,y}$	0	0	1	0
	hospital-physician cross-prospectiveness			$Pro^{cross}_{s,c,phys,hosp,y}$	1	1	1	1

Equilibrium Concept

Equilibrium, in PADSIM, is a situation in which the market for health care services clears, meaning that the quantity of services that providers want to supply equals the quantity of services that patients demand. The equilibrium condition holds, by assumption, for every combination of state, year, provider type, and coverage category:

$$Q^{supp}_{s,c,p,y} = Q^{dem}_{s,c,p,y} \ \forall\, s,c,p,y \qquad \text{eq. 1}$$

where

- s indexes states
- c indexes health insurance coverage (Medicare, Medicaid, private group, nongroup, and uninsured),
- p indexes provider types (physicians or hospitals), and
- y indexes years.

In a conventional supply-and-demand model, prices are assumed to adjust freely to maintain equilibrium. PADSIM differs from a conventional supply-and-demand model in two key ways. First, rather than there being a single price, PADSIM uses two separate concepts:

- *payment policy*: the arrangements that determine the revenues that health care providers receive from health plans and patients, and how those revenues vary with the quantity of services provided to patients
- *out-of-pocket costs*: the amounts that patients must pay to receive health care services.

Second, in PADSIM, payment policy and out-of-pocket costs do not adjust freely to maintain equilibrium; they are instead treated as exogenous inputs into the model. In contrast with a conventional supply-and-demand model that assumes prices adjust to meet demand, our approach reflects the fact that payment policy, and the level of out-of-pocket costs, are set in legislation in the Medicare and Medicaid programs, and also constrained in the small-group and nongroup health insurance markets. In large-group commercial plans, providers and plans negotiate payment policy, and plans are free to adjust out-of-pocket costs, but the current structure of PADSIM is not designed to determine private payment policy and out-of-pocket costs endogenously.

Instead of prices adjusting freely, we introduce the concept of "congestion" and use it to maintain equilibrium in PADSIM. The concept of congestion includes all nonprice factors that reduce patient demand for services and either increase, or leave unchanged, provider supply. An example of congestion is a delay between when a patient calls a physician's office and the date of the first available appointment. That type of delay is a nonprice factor—it does not relate to the out-of-pocket amount paid by the patient, nor to the payment to the provider—that may dissuade some patients from receiving services. Physicians, if they notice that patients are having to wait many days for an available appointment, may respond by expanding office hours to accommodate more appointments each day—this is an example of a supply response to congestion.

The equilibrium condition in equation 1—which simply states that supply equals demand—can be decomposed into the following:

$$e^{\ln S^{no-cong}_{s,c,p,y} + cong_{s,c,p,y}\lambda_p + \varepsilon^{supp}_{s,c,p,y}} = e^{\ln D^{no-cong}_{s,c,p,y} + cong_{s,c,p,y}\delta_p + \varepsilon^{dem}_{s,c,p,y}} \qquad \text{eq. 2}$$

where

- $S^{no-cong}_{s,c,p,y}$ is a predicted quantity of services that provider type *p* would prefer to provide to patients with coverage type *c* in state *s* in year *y* if those patients faced no congestion (hence, "no-cong")
- $D^{no-cong}_{s,c,p,y}$ is a predicted quantity of services that patients with coverage type *c* in state *s* in year *y* would prefer to receive from provider type *p* if they faced no congestion
- $cong_{s,c,p,y}$ is the level of congestion in the utilization of services of provider type *p* among patients with coverage type *c* in state *s* in year *y*
- λ_p is the elasticity of supply of provider type *p* with respect to the level of congestion (assumed to be zero or positive)
- δ_p is the elasticity of patient demand for services from provider type *p* with respect to the level of congestion (assumed to be negative)
- $\varepsilon^{supp}_{s,c,p,y}$ and $\varepsilon^{dem}_{s,c,p,y}$ are residuals.

We can then further decompose the "no-congestion" desired supply of provider type *p* as follows:

$$\ln S^{no-cong}_{s,c,p,y} = \ln N_{s,p,y} + \ln \phi_p + \ln\left(\frac{\overline{Pay}_{s,p,y}}{\overline{Pay}_{s,p,y0}}\right)\gamma_p + \overline{Pro}^{own}_{s,p,y}\eta^{own}_p + \sum_{q \neq p}\left(\overline{Pro}^{cross}_{s,p,q,y} - 1\right)\eta^{cross}_{p,q} + \ln Sshare^{no-cong}_{s,c,p,y}$$

(eq. 3)

where

- $N_{s,p,y}$ is the number of providers of type *p* in state *s* in year *y*
- ϕ_p is a supply intercept for provider type *p* (i.e., average number of services per provider)
- $\overline{Pay}_{s,p,y}$ is the weighted average real (inflation-adjusted) payment rate for provider type *p* in state *s* in year *y*
- $y0$ is a base year
- γ_p is the elasticity of supply of provider type *p* with respect to the payment rate
- $\overline{Pro}^{own}_{s,p,y}$ is the own-prospectiveness of provider type *p* (i.e., prospectiveness of revenues to provider type *p* with respect to the cost of services provided by provider type *p*) in state *s* in year *y*
- η^{own}_p is the elasticity of desired supply by provider type *p* with respect to own-prospectiveness
- $\overline{Pro}^{cross}_{s,p,q,y}$ is the cross-prospectiveness of provider types *p* and *q* (i.e., prospectiveness of revenues to provider type *q* with respect to the cost of services provided by provider type *p*) in state *s* in year *y*—note that cross-prospectiveness only affects supply if it differs from 1
- $\eta^{cross}_{p,q}$ is the elasticity of supply of services of provider type *p* with respect to cross-prospectiveness of provider types *p* and *q*

- $Sshare^{no-cong,phys}_{s,c,y}$ is the share of physician output that physicians in state *s* in year *y* would prefer to provide to patients with coverage type *c*, if no patients faced any congestion.

From the provider's perspective, the relative desirability of treating patients with different types of insurance coverage is based on the relative payment rates and prospectiveness for each coverage type. We calculate the relative financial desirability for each type of insurance coverage, and we apply a parameter that reflects the degree to which providers allocate their output based on relative financial desirability versus the relative health needs of the patient populations. We refer to this factor as "Gini power" because it affects the degree of inequality in providers' desired output to different patient types (see the appendix for more details).

Congestion

Congestion plays a key role in the equilibrium condition in PADSIM, but we do not directly observe or measure the level of congestion. Instead, we define congestion as a function of the "no-congestion" levels of supply and demand. Those no-congestion values cannot be directly observed using historical data, but they can be derived from a combination of observable historical data and a set of behavioral parameters:

$$cong^{*}_{s,c,p,y} = \mathbf{ln}\left(\frac{D^{no-cong}_{s,c,p,y}}{S^{no-cong}_{s,c,p,y}}\right)\left(\frac{1}{\lambda_p - \delta_p}\right) \qquad \text{eq. 4}$$

We can then substitute this estimate of congestion into equation 2 and rearrange:

$$\ln Q_{s,c,p,y} = \ln S^{no-cong}_{s,c,p,y}\left(1 - \frac{\lambda_p}{\lambda_p - \delta_p}\right) + \ln D^{no-cong}_{s,c,p,y}\left(\frac{\lambda_p}{\lambda_p - \delta_p}\right) + \varepsilon^{supp*}_{s,c,p,y} \qquad \text{eq. 5}$$

As is clear from this expression of the equilibrium condition, the relative elasticities of supply and demand with respect to congestion, i.e., $\left(\frac{\lambda_p}{\lambda_p - \delta_p}\right)$, is crucial. The equilibrium can be thought of as a blend of two extremes:

- If providers of type *p* are perfectly congestion-elastic—i.e., if $\left(\frac{\lambda_p}{\lambda_p - \delta_p}\right)$ is equal to 1—then the quantity of services provided will be determined by the level of patient demand. In this scenario, providers will expand output if they perceive that patients are facing access problems, regardless of the generosity of the payments they receive.
- If providers of type *p* are perfectly congestion-inelastic—i.e., if $\left(\frac{\lambda_p}{\lambda_p - \delta_p}\right)$ is 0—then the quantity of services provided will be determined by providers' preferred level of output, which is, in turn, determined by the generosity of payment. In this scenario, providers are unmoved by patients' access problems, and the level of congestion adjusts to reduce patient demand to the level providers choose to supply.

The appendix provides a step-by-step description of the computation of equilibrium quantities.

Residuals

Equation 5 can be restated as

$$\varepsilon_{s,c,p,y}^{supp*} = \ln Q_{s,c,p,y} - \left[\ln S_{s,c,p,y}^{no-cong} \left(1 - \frac{\lambda_p}{\lambda_p - \delta_p} \right) + \ln D_{s,c,p,y}^{no-cong} \left(\frac{\lambda_p}{\lambda_p - \delta_p} \right) \right] \qquad \text{eq. 6}$$

In equation 6, the residual can be calculated based on the difference between observed quantity of output and the predicted quantity of output, which is calculable by combining observed data and behavioral parameters.

3. Historical Data: Methods and Sources

PADSIM requires historical data at the state-year level on

- populations:
  - the number of individuals with each type of health insurance coverage
  - the simulated "no-congestion" demand for physician services and hospital services among individuals with each type of health insurance coverage
  - the number of physicians
  - the number of hospitals
- utilization:
  - the quantity of physician services provided to individuals with each type of health insurance coverage
  - the quantity of hospital services provided to individuals with each type of health insurance coverage
- payment policy:
  - physician payment policy (payment rates, own-provider prospectiveness, and cross-provider prospectiveness) for each type of health insurance coverage
  - hospital payment policy (payment rates, own-provider prospectiveness, and cross-provider prospectiveness) for each type of health insurance coverage.

Constructing these historical data requires significant analytical work, but they are crucial to PADSIM's operations in two ways. First, they are used to calculate utilization residuals (i.e., the natural logarithm of the ratio of the actual utilization over predicted utilization) for each combination of state, year, provider type, and source of health insurance coverage. Those residuals are carried forward and incorporated into projected utilization outcomes, and so they have a direct mechanical effect on projected spending. Second, the historical data will be used to calibrate the behavioral parameters in the model by testing how well the model fits with historical data under various behavioral assumptions. If we are testing the appropriateness of one set of behavioral assumptions versus another, we can test which set of assumptions produces residuals in the historical data that tend to be closer to zero.

Historical Data on Hospitals

The key source of data for historical measures of hospitals is the Medicare hospital cost reports (HCRs) from 1997 through 2013. We use the HCRs to measure

- the number of hospitals in each state-year ($N_{s,y}^{hosp}$)

- the quantity of hospital services (i.e., discharge-equivalents) provided to Medicare beneficiaries ($Q^{hosp}_{s,c=Mdcr,y}$), to Medicaid beneficiaries ($Q^{hosp}_{s,c=Mdcd,y}$), and to all other patients ($Q^{hosp}_{s,nonMdcrMdcd,y} = Q^{hosp}_{s,All,y} - Q^{hosp}_{s,c=Mdcr,y} - Q^{hosp}_{s,c=Mdcd,y}$)
- total revenue from providing hospital services to Medicare beneficiaries ($R^{hosp}_{s,c=Mdcr,y}$), to Medicaid beneficiaries ($R^{hosp}_{s,c=Mdcd,y}$), and to all other patients ($R^{hosp}_{s,nonMdcrMdcd,y} = R^{hosp}_{s,All,y} - R^{hosp}_{s,c=Mdcr,y} - R^{hosp}_{s,c=Mdcd,y}$).

We then allocate non-Medicare/non-Medicaid utilization and revenues to private group, nongroup, and uninsured, and calculate average payment rates as the ratio of revenues over quantity.

Historical Data on Physicians

For historical data on physicians, we combine several data sources:

- We use the Area Health Resources File (AHRF) to measure the number of nonfederal physicians whose primary activity is providing patient care ($N^{phys}_{s,y}$).
- We use data on gross state product (GSP) from the Bureau of Economic Analysis (BEA) to measure total revenues to physician offices by state and year.
- We use State Health Expenditures (SHE) data from CMS to estimate the share of revenues to physician offices from Medicare, Medicaid, and all other payers.[19]
- We use a series of publications by Stephen Zuckerman and colleagues to measure the ratio of physician payment rates in state Medicaid programs relative to Medicare.[20]
- We use national estimates of the ratio of physician payment rates for the privately insured versus Medicare from a series of reports released by the Medicare Payment Advisory Commission (MedPAC).[21]
- We use state-level estimates of the ratio of physician payment rates for the privately insured relative to the national average for the privately insured from Nguyen X. Nguyen and colleagues from the Assistant Secretary for Planning and Evaluation (ASPE).[22]

---

[19] Centers for Medicare and Medicaid Services, Office of the Actuary, National Health Statistics Group, *State Health Expenditure Accounts, 1991–2009: Converting Estimates from State of Provider to State of Residence*, Washington, D.C., December 2011.

[20] Stephen Norton and Stephen Zuckerman, "Trends in Medicaid Physician Fees, 1993–1998," *Health Affairs*, Vol. 19, No. 4, July/August 2000, pp. 222–232; Stephen Zuckerman, Joshua McFeeters, Peter Cunningham, and Len Nichols, "Changes in Medicaid Physician Fees, 1998–2003: Implications for Physician Participation," *Health Affairs*, Vol. Web Exclusive, June 23, 2004; Stephen Zuckerman, Aimee F. Williams, and Karen E. Stockley, "Trends In Medicaid Physician Fees, 2003–2008," *Health Affairs,* Vol. 28, No. 3, May/June 2009, pp. w51–w519; Stephen Zuckerman and Dana Goin, *How Much Will Medicaid Physician Fees for Primary Care Rise in 2013? Evidence from a 2012 Survey of Medicaid Physician Fees*, Kaiser Commission on Medicaid and the Uninsured, December 2012.

[21] Direct Research LLC, *Medicare Physician Payment Rates Compared to Rates Paid by the Average Private Insurer, 1999–2001*, Washington, D.C.: Medicare Payment Advisory Commission, August, 2003; Medicare Payment Advisory Commission, *Medicare Payment Policy*, Washington, D.C., March 2009, pp. 96–97; Medicare Payment Advisory Commission, *Medicare Payment Policy*, Washington, D.C., March 2012, p. 88.

Historical Data on Patients

The sources for the historical data on patients are the Current Population Survey (CPS) from 1996 through 2014 and the Medical Expenditure Panel Survey–Household Component (MEPS-HC) from 1996 through 2012. (*Patients* here refers to the entire population, i.e., all individuals who are potential users of health care services, not just the subset of individuals who actually use services.)

We use the CPS to measure the number of patients in each state-year ($N_{s,y}^{pat}$), along with demographic characteristics and type of health insurance coverage.

We use the MEPS-HC to measure

- the quantity of hospital services (outpatient, inpatient, and emergency room visits) and physician services (office-based visits with physicians, nurse practitioners and physician assistants) utilized by Medicare beneficiaries ($Q_{d,c=Mdcr,y}^{hosp}, Q_{d,c=Mdcr,y}^{phys}$), Medicaid beneficiaries ($Q_{d,c=Mdcd,y}^{phys}, Q_{d,c=Mdcd,y}^{hosp}$), patients with private group insurance ($Q_{d,c=group,y}^{hosp}, Q_{d,c=group,y}^{phys}$), patients with private nongroup insurance ($Q_{d,c=ng,y}^{hosp}, Q_{d,c=ng,y}^{phys}$), and uninsured patients ($Q_{d,c=unins,y}^{hosp}, Q_{d,c=unins,y}^{phys}$)
- the revenue from payments for hospital services (outpatient, inpatient, and emergency room visits) and physician services (office-based visits with physicians, nurse practitioners, and physician assistants) utilized by Medicare beneficiaries ($R_{d,c=Mdcr,y}^{hosp}, R_{d,c=Mdcr,y}^{phys}$), Medicaid beneficiaries ($R_{d,c=Mdcd,y}^{phys}, R_{d,c=Mdcd,y}^{hosp}$), patients with private group insurance ($R_{d,c=group,y}^{hosp}, R_{d,c=group,y}^{phys}$), patients with private nongroup insurance ($R_{d,c=ng,y}^{hosp}, R_{d,c=ng,y}^{phys}$), and uninsured patients ($R_{d,c=unins,y}^{hosp}, R_{d,c=unins,y}^{phys}$).

In the MEPS-HC, "no-congestion" demand is simulated based on the observed levels of utilization of hospital and physician services among all individuals with private group insurance or Medicare (including zeros for individuals with no utilization). Although enrollees in those sources of coverage do, in reality, face some level of congestion, their utilization patterns are useful proxies. Estimates of demand if there were no cost-sharing ("free care") are calculated based on the out-of-pocket share and the demand curve from the RAND Health Insurance Experiment. The no-congestion, free-care demand estimates are matched to individuals in the CPS based on age, sex, health status, and type of health insurance coverage. For CPS individuals, simulated no-congestion demand with cost-sharing is calculated by adjusting the no-congestion, free-care demand estimates with the average out-of-pocket share in each insurance category in the MEPS-HC, again using the RAND Health Insurance Experiment demand curve.

Population Projections

To project the patient population in each state-year, we start with historical person-level survey data from the 2013 CPS, which includes person weights that are calibrated by the U.S. Census Bureau to

[22] Nguyen X. Nguyen, Richard G. Kronick, and Steven H. Sheingold, "Comparing Physician Payment Rates Between Medicare and Private Payers in 2009," presentation at Academy Health, June 2013.

match population totals. We then create a new set of person-level weights, one weight for each projection year, by applying weight adjustment factors. Those weight adjustment factors equal the ratio of the projected population in the target year over the population in the base year, and are calculated for each combination of state, year, sex, and five-year age group (under age 5, 5 to 9, etc., 85 and up).[23] We also project changes in health insurance coverage by applying probabilistic transitions to the person-level CPS data. The transition probabilities are assigned separately for each state, year, age group (under age 25, 25 to 64, and 65 and up), and initial source of health insurance coverage, based on recent CBO projections of transitions in coverage under the ACA.[24] The demand for health care services is then projected for each combination of state, year, and source of health insurance coverage using person-level survey data, but after applying the population projection factors and the probabilistic coverage transitions.

The population totals of hospitals and physicians in each state-year are projected starting with the actual number of providers in 2013 and then applying a projected growth factor for each projection year. The projected growth factors are set equal to the projected growth in each state for the population ages 25 through 64.

Behavioral Parameters

In setting the behavioral parameters used in the model, some parameters have a strong evidence base on which to draw (e.g., physician elasticity of desired supply with respect to the payment rate), while others do not (e.g., the “Gini power” coefficients). For the proof-of-concept analyses reported here, we assigned settings to each of the behavioral parameters shown in Table 3.1, drawing where possible from published studies. In future work, RAND will explore at least three approaches to refining these parameter settings. First, for each parameter, RAND will perform a systematic review of the published literature and choose settings from the distribution of published results. Second, RAND will perform calibration tests to identify sets of parameter settings that fit relatively well with historical data on patients, payment policy, and the quantities of services supplied. Third, RAND will explore the possibility of using gaming exercises to elicit evidence on provider behavior under different simulated payment arrangements.

---

[23] Demographics Research Group, *Projections for the 50 States and D.C.*, Weldon Cooper Center for Public Service, University of Virginia, 2015.

[24] Congressional Budget Office, *Updated Estimates of the Effects of the Insurance Coverage Provisions of the Affordable Care Act*, April 2014, Washington, D.C., April 2014.

Table 3.1. Behavioral Parameters

Parameter	Notation	Parameter Name in Runsheet	Setting for Proof-of-Concept Analyses	Notes
Physician supply intercept	ϕ_{phys}	Phys_supp_intercept	10,251	Calculated from historical national data, equal to total physician RVUs divided by number of practicing physicians.
Physician elasticity of desired supply with respect to the payment rate	γ_{phys}	Phys_supp_elast_own_p	1.5	See Clemens and Gottlieb (2014), Dunn and Shapiro (2012), and Hadley and Reschovsky (2006).
Physician elasticity of desired supply with respect to own-prospectiveness	η_{phys}^{own}	Phys_supp_elast_phys_proness	–0.3	See Epstein, Begg, and McNeil (1986), and Newhouse (1993).
Physician elasticity of desired supply with respect to cross-prospectiveness of hospital revenues with respect to cost of physician services	$\eta_{phys,hosp}^{cross}$	Phys_supp_elast_hosp_proness	–0.3	See Newhouse (1993).
Patient elasticity of demand for physician services with respect to congestion	δ_{phys}	Phys_dem_elast_cong	–1	Normalized to –1.
Physician elasticity of supply with respect to congestion	λ_{phys}	Phys_supp_elast_cong	0.3	$\left(\frac{\lambda_{phys}}{\lambda_{phys}-\delta_{phys}}\right)=0.23$, meaning that output reflects a 23% weight on patients' desired utilization, and a 77% weight on physicians' desired output. Although many decades old, the following studies provide clear evidence that physician output is not completely congestion-elastic: Stewart and Enterline (1961), Enterline (1973), and Enterline et al. (1973). There are no recent studies that we are aware of that directly inform this parameter setting.
Physician "Gini power"	G_{phys}	Phys_gini_power	0.5	Physicians shift desired output shares based on relative generosity of provider payment policy, but less than proportionally
Hospital supply intercept	ϕ_{hosp}	Hosp_supp_intercept	16,767	Calculated from historical national data, equal to total hospital services divided by number of hospitals

Parameter	Notation	Parameter Name in Runsheet	Setting for Proof-of-Concept Analyses	Notes
Hospital elasticity of desired supply with respect to the payment rate	γ_{hosp}	Hosp_supp_elast_own_p	0.6	See He and Mellor (2012) and White and Yee (2013).
Hospital elasticity of desired supply with respect to own-prospectiveness	η_{hosp}^{own}	Hosp_supp_elast_hosp_proness	–0.3	See Hadley et al. (1989), Lave and Frank (1990), and Hodgkin and McGuire (1994).
Hospital elasticity of desired supply with respect to cross-prospectiveness (physician)	$\eta_{hosp,phys}^{cross}$	Hosp_supp_elast_phys_proness	–0.3	There are no well-designed studies that we are aware of that directly inform this parameter setting, and so we used the same value as in the physician setting.
Patient elasticity of demand for hospital services with respect to congestion	δ_{hosp}	Hosp_dem_elast_cong	–1	Normalized to –1.
Hospital elasticity of supply with respect to congestion	λ_{hosp}	Hosp_supp_elast_cong	0.3	$\left(\dfrac{\lambda_{hosp}}{\lambda_{hosp}-\delta_{hosp}}\right)=0.23$, meaning that output reflects a 23% weight on patients' desired utilization, and a 77% weight on hospitals' desired output. There are no well-designed studies that we are aware of that directly inform this parameter setting, and so we used the same value as in the physician setting.
Hospital "Gini power"	G_{hosp}	Hosp_gini_power	0.5	Hospitals shift desired output shares based on relative generosity of provider payment policy, but less than proportionally

4. Operating PADSIM

The software architecture of PADSIM was designed with several goals in mind:

- *Portability*: Researchers and analysts can easily make a copy of the model, run it on their local machine or on a network server, and experiment with modifications or upgrades. The one major limitation on the portability of PADSIM is that it requires that the user have a licensed copy of SAS.
- *Efficiency*: The model identifies equilibria and produces output quickly, in a matter of seconds or minutes for each run, not hours or days.
- *Accessibility*: An analyst without any specialized programming skills can define policy scenarios, run the model, and interpret the output.

To achieve those goals, PADSIM uses a "Runbox" design, meaning that all of the required input and output files for running the model are contained within a single folder. The master version of that folder is stored on RAND's DCShare1 network drive (\\dcshare1\RSR_deliv_sys_sim\PADSIM_runboxes\PADSIM_runbox_2015_09_02). A self-contained PADSIM model resides within that folder, and that folder is designed to be copied and pasted to local machines or other network drives. The Runbox uses the date of creation (e.g., "2015_09_02") to track the version of the model.

Within the Runbox are three types of files:

1. *User input*: Files that the analyst modifies and uses to specify policy scenarios and control model runs:
   - the *shell* program (e.g., "…\PADSIM--shell--2015_09_02_v2.sas"). This SAS program is the hub of the model. Within this program, the user sets the values of macro variables that specify which policy scenarios should be run, the output statistics to be generated, and the location of the Runbox. The shell program is designed to be self-explanatory, so that knowledge of SAS programming is not required (although a licensed version of SAS is required).
   - the *runsheet* file (e.g., "…\Runsheets\ PADSIM--runsheet--RAND RR--2015_09_22.csv"). This is a CSV spreadsheet file that contains user-specified policy scenarios on the columns, and user-specified parameters on the rows.
     - Each *policy scenario* represents a possible state of the world and consists of a set of projections regarding patient and provider populations, and payment policy. "Baseline" is a specific policy scenario.
     - *Parameters* are specified by the user and that controls how the model operates. Users can specify four types of parameters in the runsheet:
   - pointers to *historical data*, i.e., path and file names for CSV spreadsheet files that contain historical data on patients, providers, and utilization
   - pointers to *projection data*, i.e., CSV spreadsheet files that contain projections of patient populations and provider populations, and projections of payment policy

- *behavioral parameters*, i.e., settings that control how health care providers are assumed to behave in PADSIM
- *run control parameters*, i.e., settings that specify the SAS code containing the "behavioral engine" (i.e., SAS software that identifies equilibrium), control the output statistics that are produced, and the years that are included in projections.

2. *Input data files*: Files that PADSIM shell program uses but that the user is not required to modify:

   - *historical aggregate data* on the number of providers and patients, payment policy, and the volume of services for each combination of state, year, provider type, and types of coverage (e.g., "…\Data_and_programs\ PADSIM--histbase_state-level--2015_09_02.csv");
   - *historical patient-level data* on patient age, health insurance coverage, and utilization of health care services (e.g., "…\Data_and_programs\ patient_histyrs_2015_08_19.sas7bdat");
   - historical and projected *inflation indexes* (e.g., "…\Data_and_programs\ PADSIM--histbase_state-level--2015_09_02.csv");
   - projections of the growth in:
     - the *patient population* (e.g., "…\Data_and_programs \PADSIM--population_projections_adjfactors--2015_08_19.csv"),
     - the *number of physicians* (e.g., "…\Data_and_programs \ PADSIM--physician_projections_adjfactors--2015_08_19.csv"), and
     - the *number of hospitals* (e.g., "…\Data_and_programs \ PADSIM--hospital_projections_adjfactors--2015_08_19.csv");
   - projected *probabilities of transitioning from one type of insurance coverage to another* (e.g., "…\Data_and_programs \PADSIM--coverage_switches--baseline--2015_08_19.csv");
   - projections of changes in *provider payment policy* (e.g., "…\Data_and_programs \paypol_baseline_2015_09_22.csv");
   - the *decision engine* (e.g., "…\Data_and_programs \decision_engine_2015_08_19.sas"), which is SAS macro that
     - applies the equilibrium condition and behavioral parameters to historical data,
     - calculates residuals from the historical data by comparing the actual historical volume of services with the volume predicted using the equilibrium condition
     - applies the equilibrium condition and behavioral parameters to projected data on patients, providers, and payment policy
     - generates projected volume and spending using the projections based on the equilibrium conditions plus the residuals based on the historical data; and
   - a list of *output statistics* to be generated for each projection scenario (e.g., "…\Data_and_programs \ PADSIM--outstats--basic--2015_08_19.csv").

3. *Output data files*. These include

   – SAS datasets and CSV files containing state-year level data on historical and projected outcomes (e.g., "…\Output\Outdata\ padsim_2015_09_22_r1_baseline.sas7bdat" and "…\Output\Outdata\ padsim_2015_09_22_r1_baseline.csv")
   – a CSV file that contains summary output statistics requested by the user (e.g., "Output\PADSIM--output--RAND RR--2015_09_22--2015-09-23-17-31.csv").

5. Examples of Simulation Results

As a proof of concept, we defined three payment policy scenarios and used PADSIM to project and compare projected utilization and revenues to providers:

- *Pre-MACRA Baseline*: a scenario in which Medicare payment policy for physicians continued to be governed by the SGR (albeit with annual overrides similar to those that have occurred historically), and CMS continued to roll out the various value-based payment incentives for physicians and hospitals that were in place before the enactment of MACRA;
- *Baseline with MACRA*: a scenario in which the SGR is repealed and replaced with a new physician payment system; and
- *Repeal the Sequester:* a variation on the pre-MACRA baseline scenario in which the sequester is discontinued beginning in 2016, rather than in 2025, as is currently scheduled.

A significant amount of analytical work goes into projecting Medicare payment rates and quantifying own- and cross-provider prospectiveness under the different policy scenarios. Those payment policy projections, which are discussed in some detail in this chapter, are fed as inputs into PADSIM, along with projections of the patient population and provider supply. PADSIM then produces projections of the quantities of services provided, and total provider revenues, under the different policy scenarios. These results are meant mainly to illustrate PADSIM's operations and output. RAND is continuing to refine its projections of payment policy under these scenarios and to explore alternative settings for behavioral parameters. As a result, future analyses of MACRA will almost certainly differ from those presented here.

Pre-MACRA Baseline

In this policy scenario, we assumed that Medicare physician payment rates would grow at an annual nominal rate of 0.5 percent through 2024 and that in 2025 physician payment rates would rise by 2 percent because of the expiration of the sequester. This assumes that the SGR would continue to be overridden and that Medicare physician fees would remain flat or grow very slowly, consistent with CBO's "Alternative Fiscal Scenario" and the CMS Office of the Actuary's "Illustrative Alternative Scenario."[25] We also assumed that the prospectiveness of Medicare physician payments would increase over time because of the expanded application of the value-based payment modifier. For example, the own-prospectiveness of Medicare payments to physicians under the pre-MACRA baseline is projected to increase by 0.04, and the physician-hospital cross-prospectiveness is projected to increase by 0.09. This reflects a projected expansion in participation in alternative payment models, most notably the MSSP, to include around one-quarter of physicians and Medicare beneficiaries in 2016.

[25] Congressional Budget Office, *The Budget and Economic Outlook: 2014 to 2024*, Washington, D.C., February 2014; John D. Shatto and M. Kent Clemens, "Projected Medicare Expenditures Under Current Law, the Projected Baseline, and an Illustrative Alternative Scenario," memo, Centers for Medicare and Medicaid Services, Office of the Actuary, August 28, 2014.

Baseline with MACRA

MACRA repealed the SGR formula and established new payment methods for Medicare providers.[26] Medicare providers will choose to participate in one of two payment mechanisms starting in 2019: the Merit-based Incentive Payment System (MIPS) or Alternative Payment Model (APM). The payment mechanisms differ in payment rates and prospectiveness.

Under MACRA, payment rates are subject to annual fee updates. Annual fee updates will remain the same from April through June 2015. In the second half of 2015, rates will increase by 0.5 percent. Rates will continue to increase by 0.5 percent per year from 2016 to 2019. From 2020 to 2025, the fees will remain the same. From 2019 to 2024, qualifying APM participants will receive an annual lump sum payment equal to 5 percent of their Medicare payment in the prior year. Starting in 2026, annual fee updates will depend on providers' selection of the two payment mechanisms. Providers participating in MIPS will receive a 0.25 percent annual update. Providers participating in APM will receive a 0.75 percent annual update. In addition to the annual updates, MIPS and APM providers will be eligible for different payment adjustments based on performance.

The MIPS combines adjustments from the Physician Quality Reporting System (PQRS), Electronic Health Records (EHR) Incentive Program, and Physician Value-Based Modifier into one payment adjustment. The MIPS payment adjustments are based on providers' performance scores. Performance is measured as a composite score based on quality of care (50 percent weight in 2019, decreasing to 30 percent in 2021 and on), resource use (10 percent weight in 2019, increasing to 30 percent in 2021 and on), clinical practice improvement (15 percent weight), and meaningful use of EHRs (25 percent weight). Providers with scores below the performance threshold (the mean or median score) are penalized, while providers above the threshold receive bonuses. The maximum penalties increase incrementally from 4 percent, 5 percent, 7 percent, to 9 percent each year from 2019 to 2022; the maximum bonuses are three times the maximum penalty. The bottom 25 percent of providers receive the maximum penalty; the rest receive adjustments based on a linear scale, with scaling factors such that the adjustments are offsetting, i.e., budget-neutral. From 2019 to 2024, additional incentive payments are available to providers with scores above the 25th percentile of the range above the performance threshold; these payments are also on a linear scale based on performance scores and aggregate payments are capped at $500 million per year.

Based on our analysis of the bonus formulas, we estimated that physicians participating in MIPS, relative to traditional fee-for-service, would increase the own-prospectiveness of physician payments by 0.01, and increase the physician-hospital cross-prospectiveness by 0.08. The increase in physician-hospital cross-prospectiveness is substantial and due to the efficiency component of the value-based modifier, which will be driven in large part by hospital utilization and spending patterns.

Prospectiveness in APMs depends on shared savings and losses and capitated payments. Several different types of models may qualify for the APM payment mechanism. In this analysis, we model APMs as a mix of the MSSP, Pioneer ACO models, and patient-centered medical homes (PCMHs).

MSSP participants are eligible for shared savings or both shared savings and losses if they meet quality performance targets.[27] Track 1 is a one-sided model with up to a 50 percent shared savings rate of the savings exceeding a minimum savings rate of 2–3.9 percent (the minimum savings rate depends on the number of beneficiaries assigned to the ACO), up to a maximum payment of 10

[26] Public Law 114-10, Medicare Access and CHIP Reauthorization Act of 2015, April 16, 2015.

[27] Centers for Medicare and Medicaid Services, *Shared Savings Program*, June 2015.

percent of the benchmark expenditures. Track 2 is a two-sided model with up to a 60 percent shared savings rate of savings exceeding a minimum savings rate of 2–3.9 percent, up to a maximum payment of 15 percent of the benchmark expenditures. Shared losses are determined by a 40–60 percent shared loss rate that is adjusted based on the quality score, a minimum loss rate of 2–3.9 percent, and maximum loss sharing of 5 percent, 7.5 percent, and 10 percent of the benchmark expenditures in the first three years of the program (2012–2014). Starting in 2015, participants may choose to take on more risk in Track 3, which is a two-sided model with up to a 75 percent shared savings rate and a maximum payment of 20 percent of the benchmark expenditures, a 40–75 percent shared loss rate, and a maximum loss of 15 percent of the benchmark expenditures.

The Pioneer ACO model involves more shared risk and options for population-based payments.[28] In 2012 and 2013, the core option involved 60–70 percent two-sided sharing of savings and losses beyond 1–2 percent of target expenditures, with a 5–15 percent maximum on shared savings and losses. In 2014, participants may also elect to have 0–50 percent of their expected fee-for-service revenue as a population-based payment. Other alternative options go up to almost fully population-based payments. To estimate prospectiveness of MSSP and Pioneer ACOs, we assumed that shared savings and losses would be similar to the performance of participants in 2012–2014.

PCMHs are a model focused on the patient, primary care, and care coordination. The *Joint Principles of the Patient-Centered Medical Home* characterize a PCMH by the following principles: personal physician, physician-directed medical practice, whole-person orientation, care is coordinated and/or integrated, quality and safety are hallmarks of the medical home, enhanced access to care, and payment appropriately recognizes added value provided to patients.[29] We assumed that, in the PCMH model, 10 percent of Medicare revenues to physicians would be paid on a per capita basis.

In the MACRA scenario, we assumed that the share of physicians participating in an APM would rise to 75 percent by 2020, which is substantially higher than the 25 percent assumed in the pre-MACRA baseline scenario.

Repeal the Sequester

The “sequester” is a provision of the Budget Control Act (BCA) of 2011, which applies across-the-board spending cuts to both discretionary and mandatory spending.[30] In the Medicare program, sequestration results in a 2 percent reduction in payment rates for services provided by physicians, hospitals, and all other health care providers, as well as premium payments to Medicare Advantage plans. The sequester is scheduled to expire in 2024, at which point the 2 percent reduction in payment rates will no longer be applied. In the “repeal the sequester” policy scenario, we assumed that the 2 percent reduction in Medicare payment rates is eliminated beginning in 2016 rather than in 2025.

---

[28] Centers for Medicare and Medicaid Services, *Pioneer ACO Model*, Washington, D.C., August 2015.

[29] American Academy of Family Physicians, American Academy of Pediatrics, American College of Physicians, and American Osteopathic Association, *Joint Principles of Patient-Centered Medical Home*, March 2007.

[30] Congressional Research Service, *The Budget Control Act of 2011*, Washington, D.C., August 19, 2011.

Projections of Payment Rates and Prospectiveness

For each of these three policy scenarios, projected payment rates relative to 2013 levels are shown in Figure 5.1, and projected increases in prospectiveness relative to 2013 are shown in Figures 5.2 and 5.3. Relative to the pre-MACRA baseline, repealing the sequester results in a straightforward 2 percent increase in hospital and physician payment rates from 2016 through 2024, and has no effect on prospectiveness. Relative to the pre-MACRA baseline, MACRA results in a temporary increase in physician payment rates during the years when the 5 percent bonuses are in effect (2019–2024), but physician payment rates fall below the pre-MACRA beginning in 2025. Although MACRA did include “pay-fors” that affected payment rates for nonphysician providers, including hospitals, we did not incorporate any of those impacts on hospital payment rates in the MACRA scenario.

In general, MACRA is projected to increase prospectiveness in Medicare payments to physicians and hospitals by expanding physician participation in APMs. These increases in prospectiveness appear in physicians’ and hospitals’ own-prospectiveness, and also in hospital-physician cross-prospectiveness (i.e., the prospectiveness of hospital revenues with respect to physician costs). Somewhat counterintuitively, however, MACRA is projected to slightly decrease physician-hospital prospectiveness—that decline reflects the fact that MACRA is projected to shift substantial numbers of physicians from MIPS to a PCMH. Whereas MIPS includes a value-based payment adjustment based, in part, on hospital spending, we assumed that the PCMH model would not put physicians at risk based on hospital spending.

Simulation Results

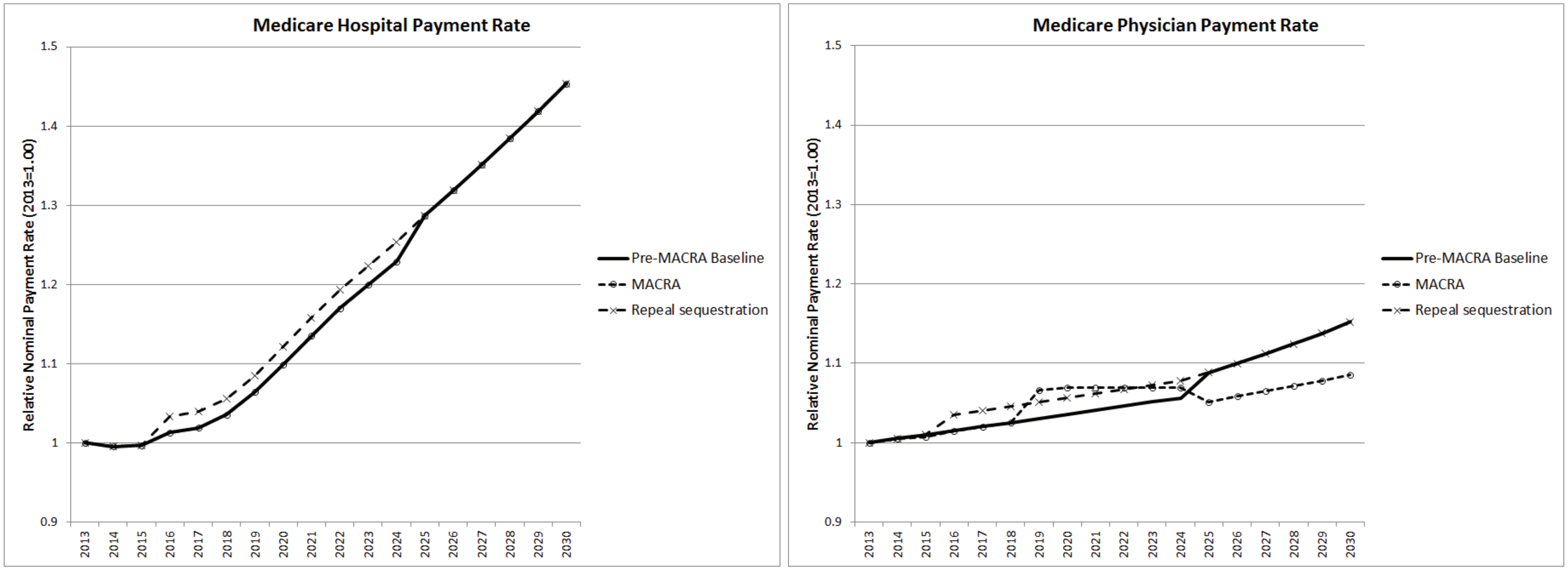
Projected spending on hospital and physician services are shown in Figures 5.4 and 5.5 for each of these three policy scenarios. The increase in revenues are partially driven by increases in payment rates and amplified by a response in volume.

In all three policy scenarios, hospital revenues continue a steep upward trajectory through 2030. Repealing the sequestration increases hospital revenues relative to the baseline, while MACRA has only a negligible impact on hospital revenues.

Both the MACRA and repeal the sequester scenarios impact physician revenues. Physician revenues under MACRA increase in 2019, partly because of the 5 percent bonus payments for APM participants, peaking at $297 billion (4.4 percent higher than revenues in the baseline). Revenues decline after 2019, partly because fee-for-service payment rates are frozen through 2025 and as more physicians shift into APMs. By 2025, physician revenues in MACRA are nearly $20 billion less than revenues in the baseline (6.6 percent lower than revenues in the baseline). From 2025 and on, physician revenues in MACRA increase at a slower rate than revenues in baseline. In the repeal sequestration scenario, physician revenues increase with the elimination of the 2 percent reduction in payment rates in 2015, and realigns with the baseline scenario in 2025 after the sequestration is scheduled to expire in the baseline. During the 2019–2024 period, MACRA is projected to substantially increase physician participation in APMs, and those payment models generally entail higher levels of prospectiveness. By itself, the increases in prospectiveness spurred by MACRA will tend to restrain Medicare utilization and spending, but those impacts are overshadowed by the additional spending resulting from the 5 percent bonuses. After 2024, Medicare physician revenues in the MACRA scenario fall further and further below the pre-MACRA baseline, which reflects the

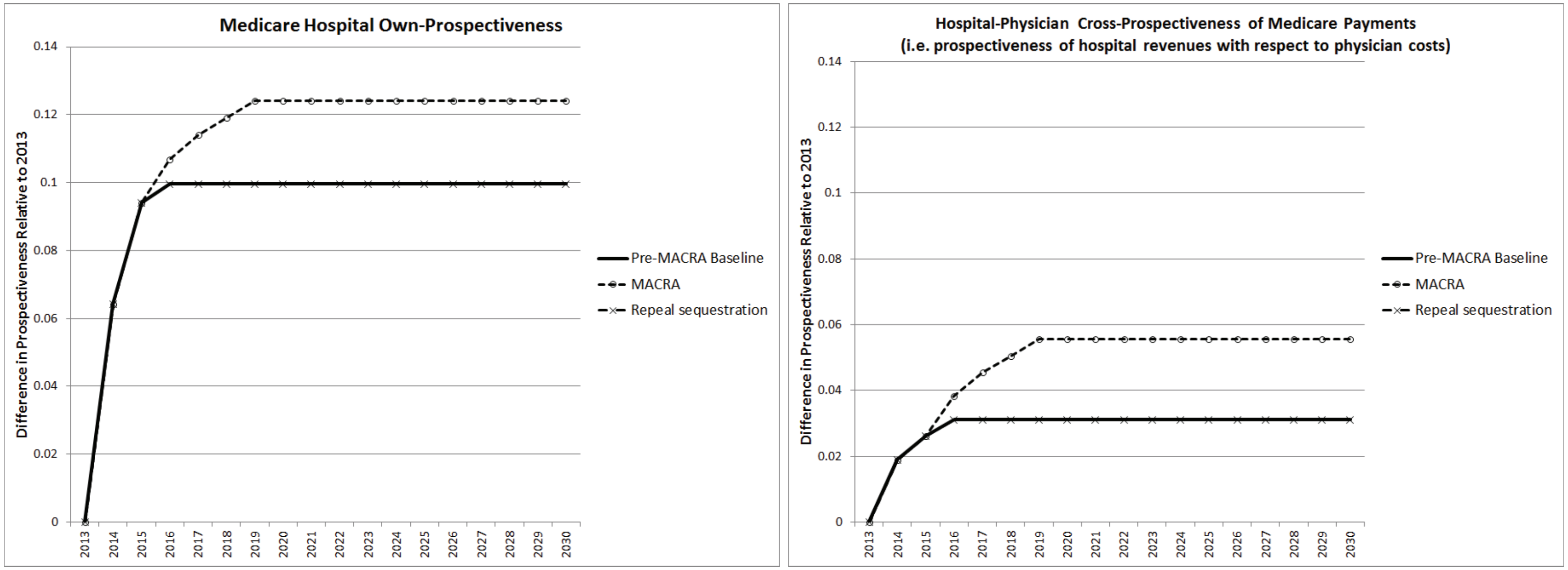
combination of increased participation in APMs, the expiration of the 5 percent bonuses, and slower growth in base payment rates.

Figure 5.1. Projected Medicare Payment Rates Relative to 2013



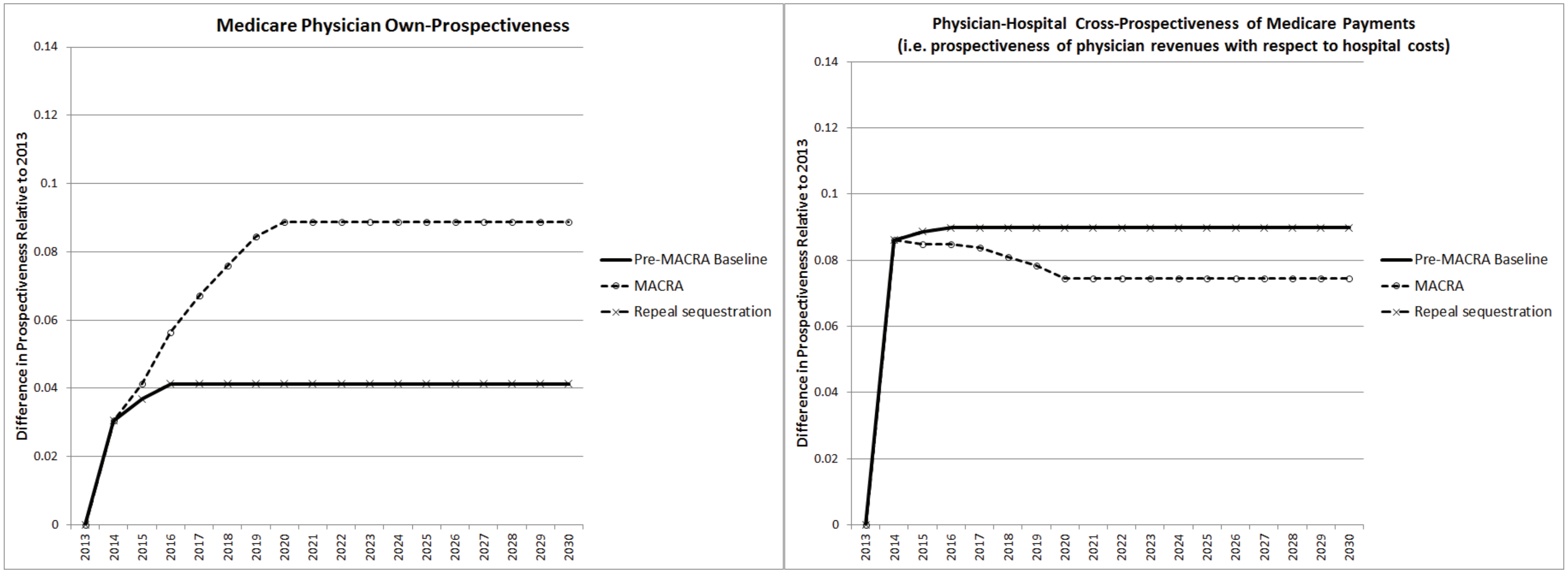
SOURCE: Authors' analysis.

Figure 5.2. Projected Increase in Prospectiveness of Medicare Payments to Hospitals



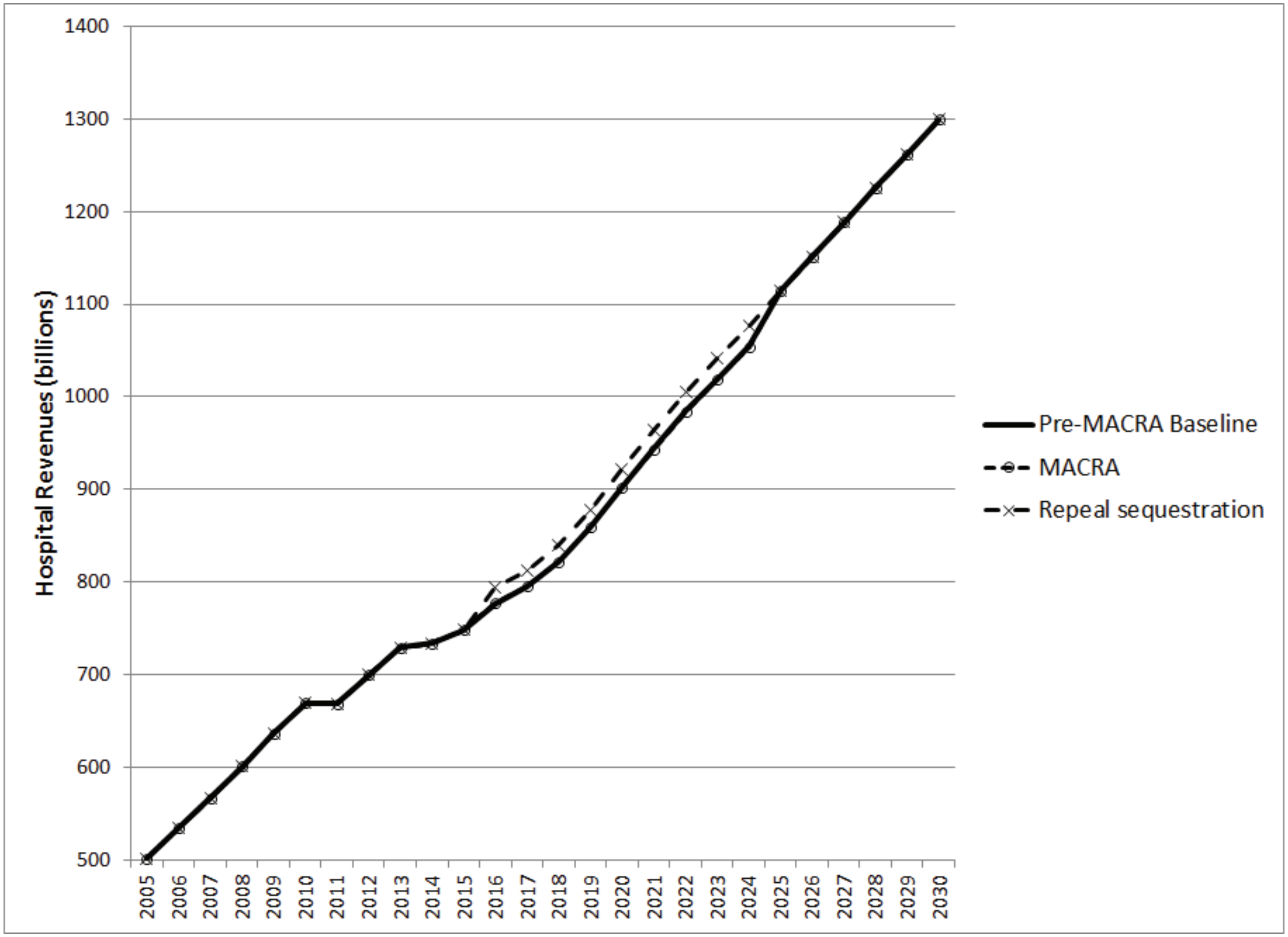
SOURCE: Authors' analysis

Figure 5.3. Projected Increase in Prospectiveness of Medicare Payments to Physicians



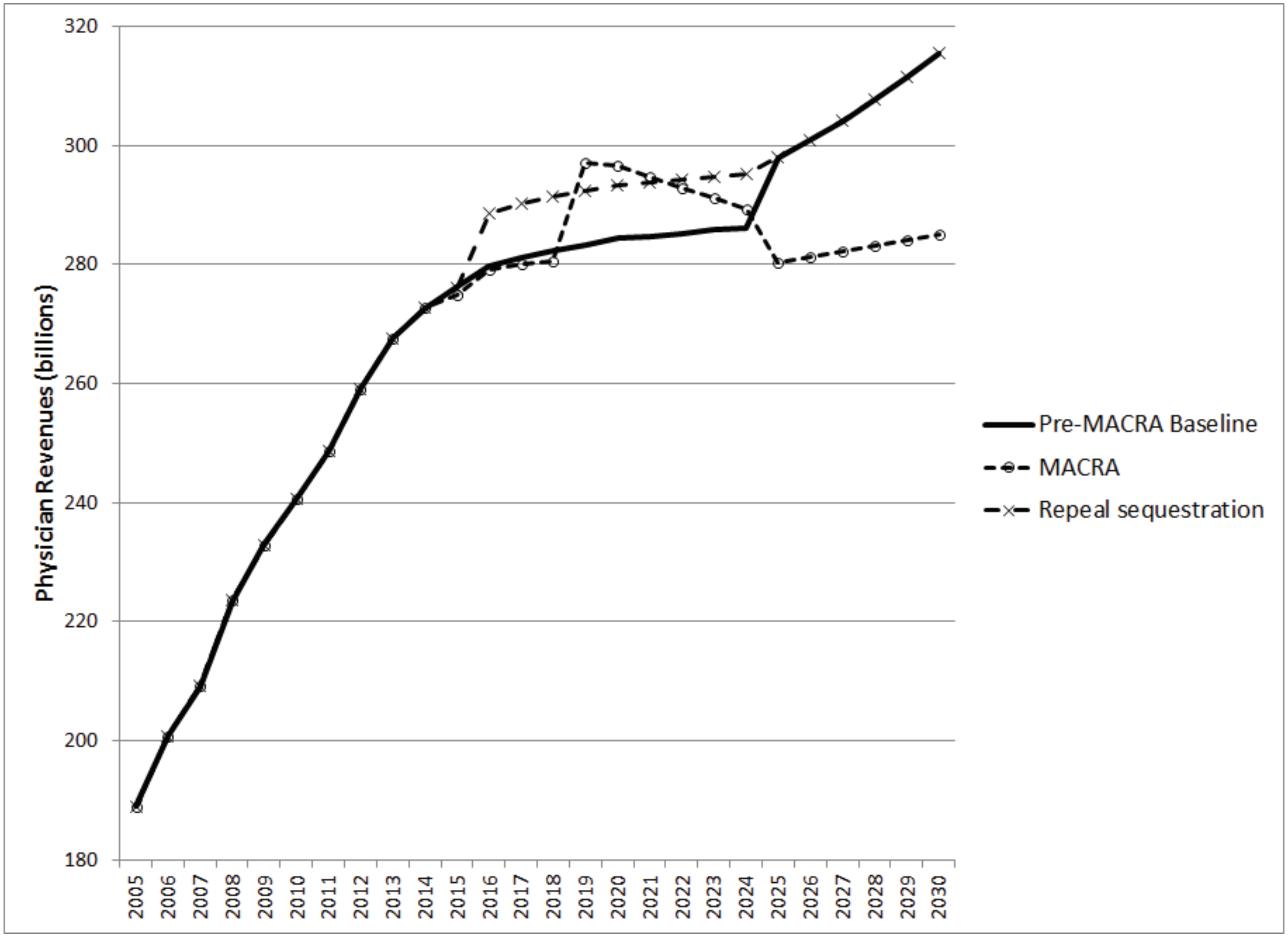
SOURCE: Authors' analysis.

Figure 5.4. Simulation Results, Hospital Revenues



SOURCE: Authors' analysis.

Figure 5.5. Simulation Results, Physician Revenues



SOURCE: Authors' analysis.

Appendix

Step-by-Step Computation of Equilibrium Quantities

To identify the equilibrium, we follow these steps:

1. *Simulate "no-congestion" patient demand*. We first simulate the "no-congestion" demand for services, for which the general formula is:

$$D^{no-cong}_{s,c,p,y} = \sum_{i \in Pop_{s,c,y}} X_i \beta_p f(OOP_c) \qquad \text{eq. 1}$$

where

- i indexes individuals
- $Pop_{s,c,y}$ is the set of individuals living in state s enrolled in coverage c in year y
- $X_{i,}$ is a vector of characteristics of individual i
- β_p is a vector that predicts utilization of service type p based on X assuming that patients face no congestion and no cost-sharing
- $f(OOP_c)$ is a function that simulates the decrease in demand from the cost-sharing that applies in coverage type c.

The data sources for calculating $X_i \beta_p$ and $f(OOP_c)$ are described in the main text (see "Historical Data on Patients").

2. *Simulate demand shares for each health insurance coverage type:*

$$Dshare^{no-cong}_{s,c,p,y} = \frac{D^{no-cong}_{s,c,p,y}}{\sum_c D^{no-cong}_{s,c,p,y}} \qquad \text{eq. 2}$$

3. *Calculate a demand-weighted average payment rate, and demand-weighted average prospectiveness:*

$$\overline{Pay}^{Dwtd}_{s,c,p,y} = \sum_c Dshare^{no-cong}_{s,c,p,y} Pay_{s,c,p,y} \qquad \text{eq. 3}$$

$$\overline{Pro}^{Dwtd}_{s,c,p,y} = \sum_c Dshare^{no-cong}_{s,c,p,y} \left(Pro^{own}_{s,c,p,y} + \sum_{q \neq p} \left(Pro^{cross}_{s,c,p,q,y} - 1 \right) \right) \qquad \text{eq. 4}$$

4. *Calculate a relative average payment rate, and a relative prospectiveness, for each coverage category:*

$$RelPay^{Dwtd}_{s,c,p,y} = \ln \left(\frac{Pay_{s,c,p,y}}{\overline{Pay}^{Dwtd}_{s,c,p,y}} \right) \qquad \text{eq. 5}$$

$$RelPro^{Dwtd}_{s,c,p,y} = Pro^{own}_{s,c,p,y} + \sum_{q \neq p} \left(Pro^{cross}_{s,c,p,q,y} - 1 \right) - \overline{Pro}^{Dwtd}_{s,c,p,y} \qquad \text{eq. 6}$$

These reflect the difference between the payment policy for coverage category *c* and the average payment policy for all coverage categories.

5. *Calculate a relative desirability of each coverage category:*

$$RelDesir_{s,c,p,y} = RelPay^{Dwtd}_{s,c,p,y} \gamma_p + RelPro^{Dwtd}_{s,c,p,y} \eta_p \qquad \text{eq. 7}$$

A relative desirability of 1 can reflect one of two situations: the first is that payment policy for coverage category *c* is equally as desirable to the provider as the average for all coverage categories, and the second is that the provider is unresponsive to payment policy (i.e. $\gamma_p = 0$ and $\eta_p = 0$). A relative desirability greater than 1 would occur of the payment policy for coverage category *c* is more generous than the average, and less than 1 if payment policy is less generous.

6. *Apply a "Gini power" factor:*

$$RelDesir^{Gpowered}_{s,c,p,y} = \left(RelDesir_{s,c,p,y} \right)^{G_p} \qquad \text{eq. 8}$$

Setting the Gini power factor, G_p, equal to 0 produces relative desirabilities for all coverage categories equal to 1—this would be appropriate if providers ignore the relative desirability of the payment policy for different coverage types when allocating their services. Setting a Gini power factor greater than 0 is appropriate if providers choose how to allocate their services based, in part, on relative desirability of the payment policy of different coverage types. This factor is referred to as the "Gini power" because it affects the degree of inequality in providers' desired output to different patient types.

7. *Calculate the share of services that providers would prefer to supply to each coverage category:*

$$Sshare^{no-cong}_{s,c,p,y} = \frac{Dshare^{no-cong}_{s,c,p,y} RelDesir^{scaled}_{s,c,p,y}}{\sum_c Dshare^{no-cong}_{s,c,p,y} RelDesir^{scaled}_{s,c,p,y}} \qquad \text{eq. 9}$$

8. *Calculate a supply-weighted average nominal payment rate, an average real payment rate, and an average prospectiveness:*

$$\overline{Pay}^{Swtd}_{s,p,y} = \sum_c Sshare^{no-cong}_{s,c,p,y} Pay_{s,c,p,y} \qquad \text{eq. 10}$$

$$\overline{Pay}^{Swtd,real}_{s,p,y} = \overline{Pay}^{Swtd}_{s,p,y} i_y \qquad \text{eq. 11}$$

$$\overline{Pro}^{Swtd}_{s,p,y} = \sum_c Sshare^{no-cong}_{s,c,p,y} \left(Pro^{own}_{s,c,p,y} + \sum_{q \neq p} \left(Pro^{cross}_{s,c,p,q,y} - 1 \right) \right) \qquad \text{eq. 12}$$

where i_y is an inflation index.

9. *Calculate a logged "no-congestion" desired level of supply to each coverage type:*

$$\ln S^{no-cong}_{s,c,p,y} = \ln N_{s,p,y} + \ln \phi_p + \ln\left(\frac{\overline{Pay}^{Swtd,real}_{s,p,y}}{\overline{Pay}^{Swtd,real}_{s,p,y0}}\right)\gamma_p + \overline{Pro}^{own}_{s,p,y}\eta^{own}_p + \sum_{q \neq p}\left(\overline{Pro}^{cross}_{s,p,q,y} - 1\right)\eta^{cross}_{p,q} + \ln Sshare^{no-cong}_{s,c,p,y}$$

(eq. 13)

and exponentiate:

$$S^{no-cong}_{s,c,p,y} = \exp\left(\ln S^{no-cong}_{s,c,p,y}\right) \qquad \text{eq. 14}$$

Or, equivalently:

$$S^{no-cong}_{s,c,p,y} = N_{s,p,y}\phi_p \left(\frac{\overline{Pay}^{Swtd,real}_{s,p,y}}{\overline{Pay}^{Swtd,real}_{s,p,y0}}\right)^{\gamma_p} \exp\left[\overline{Pro}^{own}_{s,p,y}\eta^{own}_p\right] \exp\left[\sum_{q \neq p}\left(\overline{Pro}^{cross}_{s,p,q,y} - 1\right)\eta^{cross}_{p,q}\right] Sshare^{no-cong}_{s,c,p,y}$$

(eq. 15)

10. *Calculate the logged ratio of "no-congestion" demand and supply for each coverage type:*

$$\ln XSdem_{s,c,p,y} = \ln\left(\frac{D^{no-cong}_{s,c,p,y}}{S^{no-cong}_{s,c,p,y}}\right) \qquad \text{eq. 15}$$

11. *Calculate the estimated level of congestion:*

$$cong^{*}_{s,c,p,y} = \mathbf{ln}\, XSdem_{s,c,p,y}\left(\frac{1}{\lambda_p - \delta_p}\right) \qquad \text{eq. 16}$$

12. *Calculate the predicted level of output, with congestion:*

$$\hat{Q}^{noresid}_{s,c,p,y} = D^{no-cong}_{s,c,p,y}\, \mathbf{exp}\left(cong^{*}_{s,c,p,y}\delta_p\right) \qquad \text{eq. 17}$$

Note that, by definition, we can also calculate predicted output with congestion based on supply:

$$\hat{Q}^{noresid}_{s,c,p,y} = S^{no-cong}_{s,c,p,y}\, \mathbf{exp}\left(cong^{*}_{s,c,p,y}\lambda_p\right) \qquad \text{eq. 18}$$

13. *If using historical data, calculate a residual:*

$$\varepsilon_{s,c,p,y} = \ln\left(\frac{Q_{s,c,p,y}}{\hat{Q}^{noresid}_{s,c,p,y}}\right) \qquad \text{eq. 19}$$

where $Q_{s,c,p,y}$ is the observed historical quantity.

14. *If projecting, calculate a projected quantity that includes a historical residual:*

$$\hat{Q}^{withresid}_{s,c,p,y} = \hat{Q}^{noresid}_{s,c,p,y} \exp\left(\varepsilon_{s,c,p,y0}\right) \qquad \text{eq. 20}$$

where $y0$ is the last year for which historical data are available.

References

American Academy of Family Physicians, American Academy of Pediatrics, American College of Physicians, and American Osteopathic Association, *Joint Principles of Patient-Centered Medical Home*, March 2007. As of March 3, 2016: https://www.acponline.org/running_practice/delivery_and_payment_models/pcmh/demonstrations/jointprinc_05_17.pdf

Berenson, R. A., J. H. Sunshine, D. Helms, and E. Lawton, "Why Medicare Advantage Plans Pay Hospitals Traditional Medicare Prices," *Health Affairs,* Vol. 34, No. 8, 2015, pp. 1289–1295. As of March 3, 2016: http://content.healthaffairs.org/content/34/8/1289.full.pdf

Blumberg, Linda J., Yu-Chu Shen, Len M. Nichols, Matthew Buettgens, Lisa C. Dubay, and Stacey McMorrow, *The Health Insurance Reform Simulation Model (HIRSM): Methodological Detail and Prototypical Simulation Results*, Washington, D.C.: Urban Institute, July 2003. As of March 3, 2016: http://www.urban.org/UploadedPDF/410867_HIRSM_Report.pdf

Burwell, Sylvia M., "Setting Value-Based Payment Goals—HHS Efforts to Improve U.S. Health Care," *New England Journal of Medicine*, 2015, pp. 897–899.

Centers for Medicare and Medicaid Services, *2014 Measure Information About the Per Capita Costs for All Attributed Beneficiaries Measure, Calculated for the Value-Based Payment Modifier Program*, Washington, D.C., April 2015. As of March 3, 2016: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/Downloads/2014-MIF-TPCC.pdf

Centers for Medicare and Medicaid Services, *Projections of National Health Expenditures: Methodology and Model Specification*, Washington, D.C., September 3, 2014. As of March 3, 2016: http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/Downloads/ProjectionsMethodology.pdf

Centers for Medicare and Medicaid Services, *Shared Savings Program*, Washington, D.C., June 2015, As of March 3, 2016: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/sharedsavingsprogram/index.html

Centers for Medicare and Medicaid Services, *Pioneer ACO Model*, Washington, D.C., August 2015, As of March 3, 2016: http://innovation.cms.gov/initiatives/Pioneer-aco-model/

Centers for Medicare and Medicaid Services, Office of the Actuary, National Health Statistics Group, *State Health Expenditure Accounts, 1991–2009: Converting Estimates from State of Provider to State of Residence*, December, 2011. As of March 3, 2016: http://www.cms.gov/NationalHealthExpendData/downloads/res-methodology.pdf

Clancy, G. M., and B. E. Hillner, "Physicians as Gatekeepers: The Impact of Financial Incentives," *Archives of Internal Medicine*, Vol. 149, 1989, pp. 917–920.

Clemens, Jeffrey, and Joshua D. Gottlieb, "Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?" *American Economic Review*, Vol. 104, No. 4, 2014, pp. 1320–1349.

Congressional Budget Office, *CBO's Health Insurance Simulation Model: A Technical Description*, Washington, D.C., October 2007. As of March 3, 2016:
http://www.cbo.gov/ftpdocs/87xx/doc8712/10-31-HealthInsurModel.pdf

Congressional Budget Office, Cost Estimate for the Amendment in the Nature of a Substitute for H.R. 4872, Incorporating a Proposed Manager's Amendment Made Public on March 20, 2010, Washington, D.C., March 20, 2010. As of March 3, 2016:
http://www.cbo.gov/ftpdocs/113xx/doc11379/AmendReconProp.pdf

Congressional Budget Office, *Factors Underlying the Growth in Medicare's Spending for Physicians' Services*, Washington, D.C., June, 2007. As of March 3, 2016:
http://www.cbo.gov/ftpdocs/81xx/doc8193/06-06-MedicareSpending.pdf

Congressional Budget Office, *Key Issues in Analyzing Major Health Insurance Proposals*, Washington, D.C., December 2008. As of March 3, 2016:
http://www.cbo.gov/ftpdocs/99xx/doc9924/12-18-KeyIssues.pdf

Congressional Budget Office, *The Budget and Economic Outlook: 2014 to 2024*, Washington, D.C., February 2014. As of March 3, 2016:
http://www.cbo.gov/sites/default/files/cbofiles/attachments/45010-Outlook2014_Feb.pdf

Congressional Budget Office, *Updated Estimates of the Effects of the Insurance Coverage Provisions of the Affordable Care Act, April 2014*, Washington, D.C., April 2014. As of March 3, 2016:
http://www.cbo.gov/sites/default/files/cbofiles/attachments/45231-ACA_Estimates.pdf

Congressional Research Service, *The Budget Control Act of 2011*, Washington, D.C., August 19, 2011. As of March 3, 2016:
http://fas.org/sgp/crs/misc/R41965.pdf

Cutler, David M., "The Incidence of Adverse Medical Outcomes Under Prospective Payment," *Econometrica*, Vol. 63, No. 1, January 1995, pp. 29–50.

Demographics Research Group, *Projections for the 50 States and D.C.*, Weldon Cooper Center for Public Service, University of Virginia, 2015. As of March 3, 2016:
http://www.coopercenter.org/demographics/national-population-projections

Direct Research LLC, *Medicare Physician Payment Rates Compared to Rates Paid by the Average Private Insurer, 1999–2001*, Washington, D.C.: Medicare Payment Advisory Commission, August, 2003.

Dunn, Abe, and Adam Hale Shapiro, *Physician Market Power and Medical-Care Expenditures*, Washington, D.C.: U.S. Bureau of Economic Analysis, April 26, 2012. As of March 3, 2016:
http://www.bea.gov/papers/pdf/physician_market_power_and_medical_care.pdf

Eibner, Christine, and Carter C. Price, *The Effect of the Affordable Care Act on Enrollment and Premiums, With and Without the Individual Mandate*, Santa Monica, Calif.: RAND Corporation, TR-1221-CMF, 2012. As of March 3, 2016: http://www.rand.org/pubs/technical_reports/TR1221.html

Enterline, Philip, "The Distribution of Medical Services Before and After 'Free' Medical Care—The Quebec Experience," *New England Journal of Medicine*, Vol. 289, 1973, pp. 1174–1178.

Enterline, Philip, Alison D. McDonald, and J. Corbett McDonald, "Effects of 'Free' Medical Care on Medical Practice—The Quebec Experience," *New England Journal of Medicine,* Vol. 288, No. 22, May 31, 1973, pp. 1152–1155.

Epstein, Arnold M., Colin B. Begg, and Barbara J. McNeil, "The Use of Ambulatory Testing in Prepaid and Fee-for-Service Group Practices," *New England Journal of Medicine,* Vol. 314, April 24, 1986, pp. 1089–1094.

Feder, Judith, Jack Hadley, and Stephen Zuckerman, "How Did Medicare's Prospective Payment System Affect Hospitals?" *New England Journal of Medicine,* Vol. 317, No. 14, October 1, 1987, pp. 867–873.

Girosi, Federico, Amado Cordova, Christine Eibner, Carole Roan Gresenz, Emmett Keeler, Jeanne Ringel, Jeffrey Sullivan, John Bertko, Melinda Beeuwkes Buntin, and Raffaele Vardavas, *Overview of the COMPARE Microsimulation Model*, WR-650, Santa Monica, Calif.: RAND Corporation, 2009. As of March 3, 2016: http://www.rand.org/pubs/working_papers/WR650.html

Glied, Sherry, and Nicholas Tilipman, "Simulation Modeling of Health Care Policy," *Annual Review of Public Health,* Vol. 31, No. 1, 2010, pp. 439–455.

Hadley, Jack, and James D. Reschovsky, "Medicare Fees and Physicians' Medicare Service Volume: Beneficiaries Treated and Services per Beneficiary," *International Journal of Health Care Finance and Economics,* Vol. 6, No. 2, June 2006, pp. 131–150.

Hadley, Jack, Stephen Zuckerman, and Judith Feder, "Profits and Fiscal Pressure in the Prospective Payment System: Their Impacts on Hospitals," *Inquiry*, Vol. 26, Fall 1989, pp. 354–365.

He, Daifeng, and Jennifer M. Mellor, "Do Changes in Hospital Outpatient Payments Affect the Setting of Care?" *Health Services Research*, Vol. 48. No. 5, October 2013.

He, Daifeng, and Jennifer M. Mellor, "Hospital Volume Responses to Medicare's Outpatient Prospective Payment System: Evidence from Florida," *Journal of Health Economics,* Vol. 31, 2012, pp. 730–743.

Health Care Financing Administration, Office of the Actuary, *Physician Volume & Intensity Response*, Washington, D.C., August 13, 1998. As of March 3, 2016: https://www.cms.gov/research-statistics-data-and-systems/research/actuarialstudies/downloads/physicianresponse.pdf

Hodgkin, Dominic, and Thomas G. McGuire, "Payment Levels and Hospital Response to Prospective Payment," *Journal of Health Economics,* Vol. 13, No. 1, March 1994, pp. 1–29.

Lave, Judith R., and Richard G. Frank, "Hospital Supply Response to Prospective Payment as Measured by Length of Stay," in Scheffler, Richard M and Louis F Rossiter, eds., *Advances in Health Economics and Health Services Research*, Greenwich, Conn.: JAI Press, 1990.

The Lewin Group, *The Health Benefits Simulation Model (HBSM): Methodology and Assumptions*, March 31, 2009. As of March 3, 2016: http://www.lewin.com/~/media/Lewin/Site_Sections/Publications/HSBMDocumentationMar09Rev.pdf

Medicare Payment Advisory Commission, *Hospital Acute Inpatient Services Payment System*, October 2014. As of March 3, 2016: http://medpac.gov/documents/payment-basics/hospital-acute-inpatient-services-payment-system-14.pdf

Medicare Payment Advisory Commission, *Medicare Payment Policy*, March 2012. As of March 3, 2016: http://medpac.gov/documents/Mar12_EntireReport.pdf

Medicare Payment Advisory Commission, *Medicare Payment Policy*, Washington, D.C., March, 2009. As of March 3, 2016: http://www.medpac.gov/documents/mar09_entirereport.pdf

Newhouse, Joseph P., and the Insurance Experiment Group, *Free for All? Lessons from the RAND Health Insurance Experiment*, Cambridge, Mass.: Harvard University Press, 1993.

Nguyen, Nguyen X., Richard G. Kronick, and Steven H. Sheingold, "Comparing Physician Payment Rates Between Medicare and Private Payers in 2009," presentation at Academy Health, June 2013. As of March 3, 2016: http://academyhealth.org/files/2013/tuesday/kronick.pdf

Nguyen, Xuan Nguyen, and Frederick William Derrick, "Physician Behavioral Response to a Medicare Price Reduction," *Health Services Research*, Vol. 32, No. 3, August 1997, pp. 283–298.

Norton, Stephen, and Stephen Zuckerman, "Trends in Medicaid Physician Fees, 1993–1998," *Health Affairs*, Vol. 19, No. 4, July/August 2000, pp. 222–232.

Public Law 114-10, Medicare Access and CHIP Reauthorization Act of 2015, April 16, 2015.

Shatto, John D., and M. Kent Clemens, *Projected Medicare Expenditures Under Current Law, the Projected Baseline, and an Illustrative Alternative Scenario*, August 28, 2014. As of March 3, 2016: http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/ReportsTrustFunds/Downloads/2014TRAlternativeScenario.pdf

Stewart, William H., and Philip E Enterline, "Effects of the National Health Service on Physician Utilization and Health in England and Wales," *New England Journal of Medicine*, Vol. 265, 1961, pp. 1187–1194.

White, Chapin, "Contrary to Cost-Shift Theory, Lower Medicare Hospital Payment Rates for Inpatient Care Lead to Lower Private Payment Rates," *Health Affairs,* Vol. 32, No. 5, May 2013, pp. 935–943.

White, Chapin, “Cutting Medicare Hospital Prices Leads to a Spillover Reduction in Hospital Discharges for the Nonelderly,” *Health Services Research*, Vol. 149, No. 5, 2014, pp. 1578–1595.

White, Chapin, “Medicare’s New Prospective Payment System for Skilled Nursing Facilities: Effects on Staffing and Quality of Care,” *Inquiry*, Vol. 43, No. 1, Winter 2005/2006, pp. 351–366.

White, Chapin, and Nguyen Nguyen, “How Does the Volume of Post-Acute Care Respond to Changes in the Payment Rate?” *Medicare & Medicaid Research Review*, Vol. 3, No. 1, 2011, pp. E1–E22.

White, Chapin, and Tracy Yee, “When Medicare Cuts Hospital Prices, Seniors Use Less Inpatient Care,” *Health Affairs*, Vol. 32, No. 10, October 2013, pp. 1789–1795.

White, Chapin, and Vivian Yaling Wu, “How Do Hospitals Cope with Sustained Slow Growth in Medicare Prices?” *Health Services Research,* Vol. 49, No. 1, February 2014, pp. 11–31.

Zuckerman, Stephen, Aimee F. Williams, and Karen E. Stockley, “Trends in Medicaid Physician Fees, 2003–2008,” *Health Affairs*, Vol. 28, No. 3, May/June 2009, pp. w51–w519.

Zuckerman, Stephen, and Dana Goin, *How Much Will Medicaid Physician Fees for Primary Care Rise in 2013? Evidence from a 2012 Survey of Medicaid Physician Fees*, Kaiser Commission on Medicaid and the Uninsured, December 2012. As of March 3, 2016: http://www.kff.org/medicaid/upload/8398.pdf

Zuckerman, Stephen, Joshua McFeeters, Peter Cunningham, and Len Nichols, “Changes in Medicaid Physician Fees, 1998–2003: Implications For Physician Participation,” *Health Affairs*, June 2004, pp. w4-374–w4-384.