Payment Design and the Direction of Medical Innovation

Parker Rogers*

November 2, 2026

Preliminary – please do not circulate

Abstract

Whether new technology reduces costs or expands capabilities shapes how technological progress affects spending, yet little is known about what determines the direction of innovation. Using the text of more than 411,000 U.S. medical device patents and Medicare payment reforms, I show that making health care providers bear more of the costs of the technologies they use shifts invention toward cost reduction. Across reforms spanning more than three decades, patent scores for cost reduction rise by one to three percentage points, or 0.8 to 1.5 standard deviations of the changes in unaffected fields over the same periods. Specifically, when providers bear the marginal costs of delivering care, innovation redirects toward preventing complications, shortening hospital stays, and moving treatment to lower-cost settings. When reforms only shift device costs onto providers, innovation redirects toward reducing device production costs. By contrast, when reforms change the overall returns to innovation, without changing cost exposure, the amount of patenting can change but the direction of innovation does not. Innovations that reduce the costs of delivering care are associated with shorter hospital stays without higher in-hospital mortality. Production-cost innovations are associated with lower manufacturing costs, with suggestive evidence that device prices fall after patent expiration. (JEL O31, I11, I18, L51)

*Indiana University, Kelley School of Business. AI assistance was used in data analysis and manuscript editing.

1 Introduction

Technology is the largest single source of long-run growth in health care costs (Newhouse, 1992; Chandra and Skinner, 2012), while also accounting for a large share of improvements in life expectancy (Cutler et al. 2006). Researchers have identified important factors that determine how much innovation occurs,[1] but relatively little is known about what determines the direction of innovation: whether inventors focus on expanding the capabilities of medicine or on delivering existing capabilities at lower cost. A longstanding prediction, articulated by Weisbrod (1991), is that how health care is paid for shapes the direction of medical innovation by determining whether providers benefit from adopting technologies that reduce the costs of care.

Medicare's 1983 payment reform illustrates how dramatically these incentives can change. Rather than reimbursing providers for what they spent on care, Medicare began paying a fixed price per admission. Consequently, when patients stayed an additional day in the hospital, hospital costs increased while reimbursement generally did not. New technologies subsequently enabled the rapid expansion of minimally invasive surgery, reducing recovery times and shortening hospital stays. Gallbladder removal, for example, shifted from open surgery, which typically required a week in a hospital, to laparoscopic surgery, which required only a few days (Steiner et al., 1994). The timing raises a question: did the new payment incentives help shape the development of technologies that reduced the costs of care?

Payment incentives can shape the direction of invention by changing which improvements adopters value. Inventors can lower a technology's cost or raise its quality. When adopters bear more of the costs of using a technology, they become more willing to pay for inventions that reduce those costs, increasing the relative returns to cost-reducing invention. Inventors respond by directing more effort toward cost reduction than toward quality improvement. A simple model formalizes this mechanism and distinguishes policies that change the relative returns to different types of innovation from those that primarily change the overall return to innovation.

I study these predictions primarily in the context of medical devices, which is well suited

[1] Larger markets (Acemoglu and Linn, 2004; Finkelstein, 2004; Ji and Rogers, 2024), lower entry and approval costs (Rogers, 2026), stronger patent protection (Budish et al., 2015), and public research funding (Azoulay et al., 2019).

to the analysis for two reasons. First, medical devices are frequently patented, allowing me to measure the direction of innovation using patent text. Second, Medicare reforms over more than three decades changed providers' exposure to device and care-delivery costs, creating variation in the relative returns to cost-reducing innovation.

Measuring whether inventions aim to reduce costs or improve quality is a central challenge. Existing patent-based measures, like citations, novelty, and "product-versus-process" classifications, do not directly distinguish between these two economically important dimensions. I address this gap by using a text classifier to measure three objectives of invention from each patent's title, abstract, and summary: improving quality, reducing device production costs, and reducing the cost of care through fewer complications, shorter hospital stays, or treatment in less costly settings. For each reform, I examine the dimension of invention corresponding to which costs shifted onto providers. I apply these measures to roughly 411,000 U.S. medical device patents between 1976 and 2024, allowing me to track changes in the direction of innovation within device classes over time.[2]

These measures are associated with economic outcomes. Firms that redirect their patenting toward reducing production costs exhibit lower manufacturing costs around the time of patent filing, while those that redirect toward quality do not. Specifically, a 10-percentage-point increase in a firm's production-cost score is associated with a decline of 6.6 percent in the ratio of cost of goods sold to revenue. Independent human coding provides a separate validation of the patent classifications.

To identify the effect of payment incentives on the direction of innovation, I exploit discrete changes in Medicare reimbursement policy, comparing affected and unaffected technologies. Several reforms increased providers' exposure to the costs of delivering care. Medicare's introduction of prospective payment in 1983 replaced cost-based reimbursement with fixed payments per hospital admission. The Bundled Payments for Care Improvement (BPCI) initiative introduced shared savings and losses over broader episodes of care, beginning in 2013, including hospital, physician, and post-acute services. Three additional reforms replaced cost-based reimbursement with fixed payments per episode or discharge in post-acute settings: home health agencies, inpatient rehabilitation facilities, and long-term care hospi-

[2]A large share of the variation in these measures occurs across technology fields and firms, potentially reflecting differences in both invention and patent language. Therefore, I focus on changes within fields rather than cross-sectional differences in levels and validate these changes against independent outcomes.

tals.[3] I also examine the expiration of separate reimbursement for outpatient devices, which increased providers' exposure to device costs rather than the broader costs of care. For comparison, I also study reforms affecting market size, regulatory entry costs, and taxes, which change other incentives without directly changing providers' exposure to treatment costs.

The estimates support the model's predictions. Following the introduction of prospective payment in 1983, invention aimed at reducing the costs of care increased in inpatient devices relative to those used mainly outside inpatient care, by 1.0 percentage point over 1983–90 and 3.3 percentage points over 1991–95. Nearly three decades later, bundled payment increased care-cost-reducing invention by 1.6 percentage points, while three post-acute payment reforms produced an increase of 1.1 percentage points. Relative to changes in unaffected technologies, these effects range from 0.9 standard deviations for bundled payment to 1.5 for the 1983 reform. Together, these results suggest that greater provider exposure to the costs of care shifts invention toward reducing those costs.[4]

Following the 1983 reform, hospital stays also shortened for both Medicare and privately insured patients, with no Medicare-specific change in in-hospital mortality. Although these changes cannot be attributed specifically to new inventions, they are consistent with the development and diffusion of technologies that reduce the costs of care.[5] Stays also fell more in the procedure categories whose patents shifted more toward shortening them.

Reforms that increase providers' exposure to different costs redirect invention toward reducing those particular costs. Bundled payment increases providers' exposure to the costs of delivering an episode of care, while the expiration of separate reimbursement for outpatient devices increases their exposure to device prices. Consistent with this distinction, bundled payment shifts invention toward reducing episode costs, with no corresponding in-

[3]The 1998 skilled nursing facility reform introduced a fixed, case-mix-adjusted payment per day rather than per episode. Although this increased facilities' incentives to reduce daily costs, additional days continued to generate additional payments. Therefore, I exclude this reform from the analysis of incentives to reduce the costs of an episode of care (Appendix Table A.1).

[4]The reforms require different approaches to measuring exposure: the 1983 and post-acute reforms use measures built at the technology group level, while bundled payment allows patents to be linked to affected procedures. The consistency of findings across these approaches makes it less likely that errors in any one exposure measure drive the results.

[5]This is consistent with Acemoglu and Finkelstein (2008), who find that the 1983 reform appears to have encouraged hospitals more exposed to it to adopt a range of new medical technologies, and who present suggestive evidence that it shortened their average stays. They also find faster adoption of neonatal intensive care, which is used almost exclusively by non-Medicare patients, and note that this is consistent with spillovers beyond Medicare patients. Their evidence addresses which technologies hospitals adopted, whereas this paper addresses which technologies inventors developed.

crease in production-cost-reducing invention. After separate device reimbursement expires, production-cost-reducing invention rises by 1.0 percentage point (0.8 standard deviations of changes in unaffected fields), while episode-cost-reducing invention falls by 0.9 percentage points (0.6 SD).

Reforms that primarily change the overall returns to innovation produce a different pattern. Medicare competitive bidding reduced reimbursement for durable medical equipment and the rate of invention (Ji and Rogers, 2024), while FDA deregulation increased invention (Rogers, 2026). Neither reform detectably changed the direction of innovation, nor did a device excise tax.

Payment reform can change both the types of technologies firms develop and how inventive activity is distributed across firms. About a third of the overall shift toward cost-reducing invention comes from firms changing the direction of their own invention. Another third comes from cost-focused firms shifting their inventive activity into affected technologies, with most of the remainder reflecting a larger share of invention by incumbent cost-focused firms, rather than entry by new firms.[6]

The model predicts that the same mechanism should operate when patients, rather than providers, choose and pay for medical technologies. I test this prediction using drugs that switch from prescription to over-the-counter (OTC) status.[7] These switches give patients greater control over which drugs to purchase while also exposing them more directly to drug costs. Using an analogous measure of the direction of pharmaceutical innovation, which I validate using FDA approval data, I find that the score for invention directed toward reducing patients' cost of using the drug rises by 3.6 percentage points following OTC switches, roughly one standard deviation of the changes observed among drugs that never switched.[8] These results suggest that the direction of innovation responds to cost exposure when both providers and patients choose technologies.

This paper connects the literature on directed technical change, which studies how relative

[6]The results are not driven by differences in how patents are written. The estimates remain essentially unchanged when comparing patents filed in the same year by the same firm or drafted by the same law firm.

[7]Medical devices offer limited scope for testing this prediction: patients choose only about 50 of the roughly 1,050 FDA product codes in my patent panel, and typically pay for these out of pocket, while insurers generally pay for devices selected by physicians and hospitals.

[8]Using FDA Orange Book data, I link drug patents to approved drugs. Patents linked to drugs with new dosage forms or formulations, which typically make drugs easier to take, score 1.4 standard deviations higher on the patient cost score than patents connected to approved drugs with new molecular entities.

returns shape the direction of innovation,[9] with the literature on medical innovation and health care payment. Much of the empirical literature on medical innovation examines how economic incentives affect the amount of innovation. Larger markets attract more drug and device innovation (Acemoglu and Linn, 2004; Finkelstein, 2004; Dubois et al., 2015), while regulatory approval costs and Medicare reimbursement changes affect device entry and innovation (Rogers, 2026; Ji and Rogers, 2024). Other work examines the research horizon (Budish et al., 2015) or the adoption of existing technologies. For example, Acemoglu and Finkelstein (2008) study how Medicare's 1983 payment reform affected hospitals' choices of capital and labor. This paper instead examines whether payment incentives change which technologies inventors develop.

Closest to this paper, Clemens and Rogers (2026) show that inventors focused on reducing the cost of prosthetic devices under the fixed-price procurement of the Civil War but not under the less cost-conscious procurement of World War I, suggesting that residual claimancy can affect the direction of invention.[10] I examine whether this mechanism operates more broadly across the health care system. Using eleven payment, tax, and regulatory reforms spanning more than three decades, I show that changes in providers' exposure to costs redirect invention toward reducing those costs, while reforms affecting the overall returns to innovation primarily change its amount. The distinction between device production costs and the costs of delivering care further allows me to test whether invention responds to the particular costs that providers bear. In doing so, the paper extends work on prospective payment and the delivery of care (Dafny, 2005; Cutler, 1995; Finkelstein et al., 2018) to the technologies that are subsequently developed.

The paper also contributes to a growing literature that uses patent text to measure characteristics of invention. Existing measures capture dimensions such as economic value (Kogan et al., 2017), novelty (Kelly et al., 2021), and process versus product innovation (Bena and Simintzi, 2025). I instead use patent text to distinguish inventions aimed at reducing costs from those aimed at improving quality, a dimension that is empirically distinct from these existing measures. I validate the classifications against firms' manufacturing costs, hospital stays across procedure categories, and independent human coding.

[9] Acemoglu (2002); Hanlon (2015); Aghion et al. (2016).

[10] Residual claimancy refers to the extent to which the party choosing a technology retains the savings from reducing costs. See Laffont and Tirole (1993); Rogerson (1994).

The paper proceeds as follows. Section 2 describes the Medicare payment reforms. Section 3 develops the model. Section 4 constructs and validates the measures of cost-reducing and quality-expanding innovation, and Section 5 presents the empirical strategy. Section 6 estimates the effects of payment reform on the direction and amount of innovation. Section 7 examines the economic consequences of cost-reducing innovation. Section 8 examines whether the mechanism extends to patients as decision-makers, and Section 9 discusses the policy implications and concludes.

2 Background

Medicare reimburses providers through several payment arrangements that differ in the extent to which providers bear the marginal costs of care. Some arrangements issue additional reimbursement when providers use more expensive devices or deliver more care, while others offer only fixed reimbursement, incentivizing providers to reduce treatment costs. Over time, Medicare has introduced several reforms that changed these incentives. I describe the main reforms below, with additional institutional details in Appendix A.

Before 1983, Medicare generally reimbursed hospitals based on their reported costs. Hospitals, therefore, had relatively weak incentives to reduce the cost of treating an additional patient. In 1983, Medicare introduced the Inpatient Prospective Payment System, replacing cost-based reimbursement with payments determined primarily by a patient's diagnosis-related group (DRG). Hospitals could retain the savings from less costly treatment but generally received no additional reimbursement when treatment became more expensive.[11] The 2000 Outpatient Prospective Payment System (OPPS) introduced a similar arrangement for hospital outpatient care.

More recent reforms extended these incentives beyond the hospital stay. Under the 1983 prospective payment system, hospitals benefited from reduced admission costs but generally had no direct financial stake in the costs of care delivered after discharge. Bundled payment expanded these incentives by holding participating providers financially accountable for spending over a broader episode of care. Under the Bundled Payments for Care Improvement (BPCI) initiative, which began in 2013, Medicare generally continued to reim-

[11] This applies to operating costs. Capital costs continued to be reimbursed on a cost basis until capital prospective payment was phased in beginning in fiscal year 1992.

burse individual services but compared total episode spending, including hospital, physician, and post-acute care, with a target price.[12] Participating providers could receive a share of savings below the target and were responsible for some excess spending. These arrangements increased the returns to technologies that reduce costs throughout an episode of care, including complications, readmissions, and post-acute care.

Medicare also introduced prospective payment in several post-acute settings. Between 2000 and 2002, it moved home health agencies, inpatient rehabilitation facilities, and long-term care hospitals from cost-based reimbursement to fixed payments per episode or discharge, which increased each facility's financial return to reducing the cost of its care.[13]

Not all devices are paid for entirely through a fixed payment. Medicare sometimes provides additional reimbursement for new technologies, including temporary pass-through payments for certain outpatient devices and new-technology add-on payments for qualifying inpatient technologies. These payments reduce the costs that hospitals bear when adopting eligible technologies. When separate reimbursement expires, hospitals generally must cover the device's price from their existing payments, increasing their incentive to use less expensive devices. Unlike reforms that increase exposure to the overall costs of care, the expiration of separate device reimbursement increases the value providers place on less expensive devices.

Other policies affect the returns to medical innovation without changing providers' exposure to treatment costs. Medicare's competitive bidding program for durable medical equipment, prosthetics, orthotics, and supplies (DMEPOS), introduced in 2011, reduced reimbursement for certain types of equipment. FDA reclassification reduced regulatory requirements for some device types, while the medical-device excise tax taxed manufacturers' sales. These policies can change the profitability of developing new devices without directly changing the share of treatment costs providers bear.

Finally, patients generally do not choose the specific medical devices used in their treatment, and their out-of-pocket payments often depend only indirectly on the devices providers select. Prescription drugs offer a setting in which the patient can become both the buyer and decision-maker. When a drug switches from prescription to OTC status, patients gain

[12] BPCI's Model 4 instead paid a single prospective amount for the hospital stay and inpatient physician services, and its Model 3 covered post-acute care only.

[13] Skilled nursing facilities moved to prospective payment in 1998, but their per-day rate does not reward shorter stays or fewer complications (Appendix Table A.1).

greater control over which drugs to purchase and typically become more directly exposed to drug prices. I examine these switches separately in Section 8.

3 Conceptual Framework

Payment systems can influence the direction of medical innovation by changing characteristics that providers value in the technologies they adopt. In this simple model, a device manufacturer can invest in improving the quality or reducing the costs of a technology. The returns to these investments depend, in part, on how much providers are willing to pay for these advancements. When providers bear more of the costs of care, they place greater value on technologies that reduce those costs, raising the returns to cost-reducing invention.

The key policy variable is how much of the costs of care providers bear. With cost-based reimbursement, insurers reimburse providers for allowable expenditures, so reducing those expenditures also reduces reimbursement. With prospective payment, providers receive a fixed payment, allowing them to benefit financially from reductions in the costs of care. For example, a technology that shortens hospital stays becomes more valuable to a hospital when the hospital can capture the savings from discharging patients earlier. Manufacturers, anticipating this demand, have greater incentives to develop technologies that reduce these costs.[14] A formal model incorporating this mechanism is provided in Appendix B.

The 1983 reform changed how hospitals were paid for inpatient care, but left payments for non-hospital services unaffected. Therefore, the reform increased the incentive to develop cost-reducing technologies used in inpatient care relative to those used in other settings. More generally, the model predicts that changes in the direction of invention will be larger for technologies used in settings affected by a reform. This prediction motivates the comparisons between affected and unaffected technologies in Section 5 (Appendix B).

Payment reforms also differ in which type of costs they incentivize providers to reduce. Prospective and bundled payments increase providers' exposure to the costs of delivering care, including hospital stays, complications, and subsequent treatment. These arrange-

[14]This mechanism builds on the argument of Weisbrod (1991) that reimbursement systems can shape the direction of medical innovation and is related to Baumgardner (1991)'s analysis of insurance and technical change.

ments increase the value of technologies that reduce the resources needed to treat a patient. By contrast, the expiration of separate reimbursement for an outpatient device increases providers' exposure specifically to the price of the affected device. Once hospitals bear the cost of purchasing a device, manufacturers that reduce production costs can attract more customers by lowering prices while retaining some of the cost savings as profit. Therefore, the types of costs affected by a reform can influence which types of cost-reducing inventions become more valuable.

Other policies primarily affect the overall returns to innovation rather than how much of the costs of care providers bear. For example, market expansions, deregulation, or tax changes can increase or decrease the amount firms invest in R&D without necessarily changing the direction of innovation, provided they do not affect the relative returns to cost reduction and quality improvement and the cost of research responds similarly on both margins (Appendix B).

While these predictions focus on providers, they also apply when patients choose and pay for medical products themselves, as in the case of OTC drugs. Table 1 summarizes the predictions and corresponding estimates.

4 Data and Measurement

4.1 Data and Sample Construction

The analysis draws primarily on U.S. medical-device patents, supplemented by data on device prices, firm manufacturing costs, and patient outcomes. I obtain 411,651 patents granted between 1976 and 2024 from the USPTO and PatentsView, identifying medical-device patents within the A61 patent classification (Appendix C). Each patent includes its filing year, technology classification, and textual description. Throughout the analysis, I date patents by filing year rather than grant year, since filing dates more closely reflect when inventions were developed. I also obtain patent citations and estimated economic values from PatentsView and Kogan et al. (2017).

I use two approaches to measure patent exposure to the reforms. When the affected products or procedures can be identified directly, I link patents to them. Specifically, for bundled

payments, pass-through payments, the device tax, and OTC switches, I construct text-based crosswalks linking patents to FDA product codes, Medicare Healthcare Common Procedure Coding System (HCPCS) codes, and FDA drug approvals. FDA product codes classify devices by type, while HCPCS codes identify products and services for reimbursement. The crosswalks use text embeddings to measure the similarity between patent descriptions and the corresponding product or procedure descriptions. For FDA reclassification and competitive bidding, I use the patent links developed in my earlier studies (Rogers, 2026; Ji and Rogers, 2024).

By contrast, for reforms that change payment across entire settings of care, I identify affected technologies using main groups of the Cooperative Patent Classification (CPC) system, which classifies patents by technology. This approach applies to the 1983 inpatient reform and the post-acute reforms. Appendix Table A.2 lists the affected technology groups. Appendix C describes the matching procedures, evaluates the accuracy of the HCPCS crosswalk, and validates the drug crosswalk against patents listed in FDA's Orange Book.

Because the reforms affect different technologies at different times, I construct a separate estimation sample for each reform. For the 1983 inpatient reform, I compare 27 CPC main groups of technologies used primarily in inpatient care with 30 groups used mainly outside inpatient care, excluding dental devices.[15] The sample contains 32,328 patents filed between 1977 and 1995. I manually classify these groups using their official CPC titles.[16] Medicare claims support this classification: among groups linked primarily to billed procedures, inpatient technologies account for a substantially higher share of inpatient procedures than the comparison technologies (54 versus 8 percent on average in 2010–2012). A continuous measure of inpatient exposure constructed using a language model produces similar estimates (Appendix C).

For bundled payment, the sample contains 74,780 patents filed between 2005 and 2020, including 10,326 linked to procedures covered by BPCI. For the three post-acute reforms, I compare 26 CPC main groups containing supplies and equipment covered by the new payment systems with 49 acute-care technology groups matched on their pre-reform charac-

[15]Including dental devices in the comparison group gives an estimate of 2.69 percentage points (s.e. 1.24) for 1991–95, compared with 3.31 in the baseline.

[16]Appendix Table A.2 lists the groups in each category.

teristics (Appendix Table A.3).[17]

For the expiration of outpatient pass-through payments, I link each patent to its closest billing code and compare 11,910 patents associated with 80 device categories that received pass-through payments with patents linked to 6,888 codes that never received them.[18]

The remaining analyses use panels of device product codes and equipment categories, while the OTC analysis uses a panel of drug molecules. Appendix C provides the sample definitions and construction details.

I supplement the patent data with several datasets containing economic outcomes that I use to validate the measures of innovation. I obtain device prices from ECRI's hospital procurement data (2012–2022), which record transactions between hospitals and device manufacturers at the SKU level across the U.S. I measure manufacturing costs using CRSP-Compustat financial data and link patents to publicly traded firms using the patent-level match from Kogan et al. (2017).[19] Patient outcomes are derived from the National Hospital Discharge Survey (1979–2007), which includes data on inpatient procedures, length of stay, and discharge status. These datasets and their construction are described in Appendix C.

4.2 Measuring the Direction of Invention

The central measurement challenge is distinguishing inventions intended to reduce costs from those intended to improve quality. Existing measures of innovation do not directly capture these important objectives, especially at the level of detail required for this analysis.[20] In response, I construct measures of the direction of invention using a trained text classifier that identifies whether inventions target different forms of cost reduction or quality improvements, based on each patent's title, abstract, and summary of invention.

[17] Including dental technologies in the comparison pool and re-matching gives an estimate of 1.34 percentage points (s.e. 0.47), compared with 1.11 in the baseline.

[18] Six additional treated categories, each linked to a single patent, and approximately 2,000 comparison codes observed in only one year of a stack are absorbed by the fixed effects. About half of the comparison codes were already billed in 2000. The estimates are unchanged when I restrict the comparison group to codes not yet billed in 2000 (Appendix C).

[19] The match assigns patents granted to publicly traded firms to their CRSP identifiers. The sample includes 1,820 firms holding at least one medical-device patent, of which 1,564 have cost of goods sold and revenue data in Compustat between 1990 and 2023.

[20] The most closely related measures, the product and process innovation classifications of Bena and Simintzi (2025), are conceptually distinct from my measure and nearly uncorrelated with it ($\kappa = -0.01$; Appendix C).

I distinguish between inventions that reduce device production costs and those that reduce the costs of delivering care. Production-cost inventions include improvements in manufacturing methods and materials that make devices less expensive to produce. Inventions that reduce the costs of care instead prevent complications, shorten hospital stays, or allow treatment in less costly settings.[21] I measure these objectives separately because payment rules differ in which costs providers bear. A fixed payment per admission rewards innovations that shorten hospital stays or prevent complications, but not those that avoid the admission altogether. Bundled payments reward all three channels, while the expiration of separate reimbursement for a device increases the incentive to reduce device prices. I therefore construct an episode-cost measure by averaging the channels relevant to each reform: complications and length of stay for payments fixed per admission or stay, and all three channels for bundled and episode payments.[22]

The classifier assigns each patent a separate score between zero and one for each objective. Because the scores are measured independently, a patent can contribute to both cost reduction and quality improvement. I use the continuous scores rather than assigning patents to mutually exclusive categories.[23] Appendix C describes the classifier, training data, and the prompt used to generate the training examples. Table 2 illustrates the distinction between the different measures using two patents in each direction.

[21]These channels do not capture every way an invention can reduce the cost of care. I also measure labor saving, including shorter procedures and reduced staffing requirements. Patents describing these improvements score at approximately the 66th percentile of the labor-saving measure. I exclude labor saving from the episode-cost score because I do not observe a corresponding outcome with which to validate it, and because it is the least accurately classified objective (an area under the receiver operating characteristic curve, AUC, of 0.77 against held-out labels). I report it separately in Section 6.3. Inventions that eliminate diagnostic tests are not classified separately. Patents describing such technologies tend to score highly on changes in the site of care, as with point-of-care testing, but not on complications or length of stay.

[22]The results are similar when I weight channels by classification accuracy, standardize them before averaging, or use their first principal component. Across all designs, these alternatives produce estimates between 0.9 and 1.1 times the equal-weight estimate.

[23]The classifiers give greater weight to less common objectives during training, so the scores are not calibrated to their prevalence among patents. A one-percentage-point increase in a score therefore does not imply a one-percentage-point increase in the share of patents pursuing that objective. I report effects in percentage points of the score and compare them with the typical changes observed among technologies unaffected by the reforms (Section 7.4).

4.3 Measurement Validation

I evaluate the classifier by comparing its scores with independent human classifications of device patents. Using the average ratings of two human coders, the production-cost measure achieves an AUC of 0.84.[24] The corresponding AUCs are 0.78 for quality improvement and 0.76 for complication reduction, although agreement between the human coders is lower for these objectives.[25]

I re-estimate each design using three alternative text classifiers that use the same training labels, and find the results are broadly consistent. Appendix C reports the estimates for each classifier. I also use the 5,460 patents in the training sample, which includes patents filed before and after 1983, to test whether classification errors change differently in technologies affected by each reform. Controlling for each patent's training label, I find no significant differential change in classifier scores following any reform.

I also validate the measures against changes in manufacturing costs and hospital length of stay (Figure 1). Among publicly traded device manufacturers, a 10-percentage-point increase in the production-cost score is associated with a 6.6 percent lower ratio of cost of goods sold to revenue around the time of patent filing. Higher quality-improvement scores are associated with a higher ratio, although the estimate is not statistically significant. For the stay-reducing measure, I compare changes in patenting and length of stay across 67 inpatient procedure categories. A one-standard-deviation larger increase in stay-reducing invention after 1983 is associated with a 5 percent larger decline in length of stay during 1984–1990 and an 11 percent larger decline during 1991–1995. A relationship of similar magnitude is also present before the reform. These relationships support the economic interpretation of the patent measures, although neither establishes that the inventions caused the observed

[24] An AUC of 0.84 means that, when comparing a patent classified by the human coders as targeting production costs with one classified as not targeting them, the classifier assigns the former a higher score 84 percent of the time. The validation sample consists of device patents excluded from training, with patents receiving high classifier scores oversampled to ensure sufficient positive examples for each objective. A research assistant coded 200 patents, and I independently coded 150 of them, selected by strata of the predicted production-cost score. I did not observe patent numbers, the assistant's ratings, or the classifier's scores. Each coder rated each objective from zero to one, and I classify a patent as targeting an objective when the average rating is at least 0.5. Reweighting the sample to represent the population of device patents gives a production-cost AUC of 0.74. Appendix C reports results for each coder separately.

[25] The length-of-stay and site-of-care measures have not been evaluated against human classifications. Against held-out labels generated by the labeling model, they achieve AUCs of 0.95 and 0.84, respectively (Appendix Table A.4). The length-of-stay measure also achieves an AUC of 0.90 against an independent set of 146 patents classified by a frontier language model (Anthropic's Claude Opus).

changes in costs or length of stay.

4.4 Summary Statistics

Table 3 reports summary statistics for the universe of device patents and for each reform's estimation sample: the 1983 reform, bundled payment, the post-acute reforms, pass-through expiration, and the pharmaceutical patents used in Section 8. Across 411,651 device patents, the average production-cost score is 0.21, compared with 0.50 for quality improvement and 0.32 for the episode-cost score. The production-cost score is right-skewed, with a median of 0.18 and a 90th percentile of 0.40. Most of the variation in the production-cost score occurs within technology classes and years: 84 percent within CPC-subclass-by-year cells and 71 percent within the narrower main-group-by-year cells.

Exposed technologies differ from comparison technologies before the reforms. In the 1983 sample, inpatient technologies had higher episode-cost scores before the reform than comparison technologies (0.38 vs. 0.24). By contrast, patents exposed to bundled payment had slightly lower episode-cost scores before 2011 than unexposed patents (0.36 vs. 0.38). I classify a patent as exposed to bundled payment if at least one of its five closest procedure codes in the crosswalk is covered by BPCI, restricting links to those with a similarity score of at least 0.45. Among exposed patents, BPCI-covered codes account for 42 percent of qualifying links on average. Appendix C reports estimates using this share as a continuous measure of exposure. Production-cost scores also decline over time, from an average of 0.30 for patents issued in the 1980s to about 0.20 for those issued in the 2010s.

Appendix C provides additional summary statistics, variable definitions, and details on the construction of each estimation sample.

5 Empirical Strategy

I estimate the effects of the Medicare reforms using difference-in-differences and event-study designs that compare changes in the direction of invention across technologies with different exposure to each reform. The reforms affect different products and services, so the units of observation and measures of exposure vary across analyses. The common approach is to compare inventions in more- and less-exposed technologies before and after each reform.

Appendix Table A.5 summarizes the identifying variation, comparison group, specification, and inference for each design.

5.1 Estimation

Most analyses use patents as the unit of observation and the continuous classifier scores described in Section 4 as outcomes. For each reform, I estimate

$$y_{ift} = \beta \left(D_i \times \text{Post}_t \right) + \gamma D_i + \alpha_f + \delta_t + \varepsilon_{ift}, \tag{1}$$

where y_{ift} is the score of patent i, filed in year t and belonging to technology f, for the cost-reducing objective affected by the reform. D_i indicates whether the patent is associated with an affected technology, procedure, device category, or product, while Post_t indicates years after the reform. The terms α_f and δ_t are technology and filing-year fixed effects. The coefficient β measures the change in the direction of invention among affected patents relative to comparison patents.

For the 1983 reform, I estimate separate effects for 1983–1990 and 1991–1995. For pass-through payments, I estimate separate effects during the period of separate reimbursement and after its expiration, reporting the difference between them. Appendix C reports alternative specifications for the 1983 and bundled-payment reforms that replace the binary treatment indicator D_i with a continuous measure of exposure, E_i, which find consistent results.

I also estimate event-study specifications to examine how invention changes before and after each reform:

$$y_{ift} = \sum_{k \neq -1} \beta_k \left(D_i \times \mathbf{1}\{t - t_0 = k\} \right) + \gamma D_i + \alpha_f + \delta_t + \varepsilon_{ift}, \tag{2}$$

where t_0 is the reform year and k denotes years relative to the reform. I omit the year immediately preceding the reform, so each coefficient measures the differential change relative to that year. Figure 2 instead re-centers each series on the average of the six pre-reform years, so the plotted path does not depend on a single reference year. The pre-reform coefficients allow me to examine whether invention was already changing differently across more- and

less-exposed technologies.

For reforms implemented at different times across technologies, including the post-acute prospective-payment reforms and the expiration of outpatient device pass-through payments, I use stacked difference-in-differences specifications following Cengiz et al. (2019). Each event-specific panel includes the affected technologies and their eligible controls. I then stack these panels and include event-by-technology and event-by-year fixed effects. This avoids the comparisons between already-treated and newly treated technologies that can bias conventional two-way fixed-effects estimates when treatment effects differ across cohorts (Goodman-Bacon, 2021; Sun and Abraham, 2021).

5.2 Treatment Exposure, Dependent Variables, and Comparison Groups

The primary outcome for each reform reflects the costs most directly affected by its payment arrangement. For the 1983 inpatient prospective payment reform, I focus on inventions aimed at reducing complications and shortening hospital stays. These are the costs of an inpatient admission that a payment per admission made the hospital bear. The length-of-stay measure includes earlier discharge, so it also captures inventions that let a hospital discharge a patient sooner to a less costly setting, which the 1983 rules rewarded because post-acute care was paid separately.[26] I do not include inventions that move treatment out of the hospital altogether, which comprise the site-of-care measure, because the 1983 rules gave the hospital little reason to favor these types of inventions: an avoided admission forfeited its DRG payment, and hospital outpatient care remained largely cost-based until 2000.[27]

Bundled payment makes one provider responsible for spending across hospital, physician, and post-acute care, so I include inventions that move treatment to less costly settings in its episode-cost measure. For the post-acute reforms, I instead define the outcome according to the costs each provider bears. Inpatient rehabilitation facilities and long-term care hospitals are paid per discharge, rewarding shorter stays and fewer complications. Home health agen-

[26]Until October 1998, a hospital received the full DRG payment when it discharged a patient early to a skilled nursing facility or home health agency; the Balanced Budget Act of 1997 then reduced payment for early discharges in ten DRGs.

[27]Invention aimed at that channel does not rise after 1983, and including it lowers the 1991–1995 estimate from 3.3 to 2.1 percentage points (standard error 1.1), which is marginally significant (Appendix Table A.6).

cies are paid per 60-day episode, rewarding fewer visits. Skilled nursing facilities are paid per day, so I do not classify them as exposed. Using all three episode-cost channels for every reform produces similar results (1.16 versus 1.11 percentage points; Appendix Table A.6). The expiration of separate reimbursement for outpatient devices instead increases providers' exposure to device purchase prices, so I examine production-cost invention. Appendix A describes the payment rules.

For the 1983 reform, I compare patents in 27 CPC main groups of technologies used mainly in inpatient care with patents in 30 groups used mainly outside inpatient care (Section 4). For BPCI, treatment is defined at the patent level using links to covered procedures. I compare these patents with those linked to major surgical procedures not covered by BPCI.[28]

For the three post-acute reforms, I compare 26 CPC main groups containing affected supplies and equipment with 49 acute-care groups matched on pre-reform patent scores and patenting volume. Using the broader pool of comparison technologies gives an estimate of 0.98 percentage points (s.e. 0.50; Appendix A). For outpatient pass-through payments, I stack device categories by the year separate reimbursement began, using billing codes that never received pass-through payments as comparisons. The estimate compares production-cost scores after separate reimbursement expires with those during the period of separate payment.

The remaining reforms change the overall returns to invention without changing which costs providers bear: FDA reclassification, Medicare competitive bidding for durable medical equipment, and the medical-device excise tax. For FDA reclassification and competitive bidding, I use the treated units, treatment dates, comparison groups, patent links, and stacked estimators from Rogers (2026) and Ji and Rogers (2024). I replace patent counts with the average production-cost score of patents in each unit-year, weighting observations by their patent counts. I retain the preferred matched specifications, but use cluster-robust standard errors or a wild cluster bootstrap rather than Conley–Taber intervals, which require observing every unit in every year. Appendix C describes the corresponding product-code

[28]The comparison procedures are paid as comprehensive outpatient or inpatient-only services, the payment classifications of 105 of the 111 BPCI episode procedures. Using all procedure-linked patents as comparisons gives an estimate of 1.29 percentage points (s.e. 0.44). Because treatment is measured using patent text, I also test whether the share of patents linked to bundled procedures changes after the reform. I find no significant increase after 2011, including among technologies with greater pre-reform exposure (Appendix C).

and technology-group panels. Appendix Table A.1 lists related reforms I considered but excluded.

5.3 Reform Timing

I date reforms to the earliest announcement that made the change in payment incentives sufficiently clear for manufacturers to respond. Because patents are dated by filing year, invention may respond before a reform is implemented.[29]

For the inpatient reform, I use 1983 and estimate separate effects for 1983–1990 and 1991–1995. The latter period follows a strengthening of hospitals' incentives to reduce costs, as Medicare margins turned negative and managed care limited hospitals' ability to offset payment cuts through higher private prices (Cutler, 1998). Section 6.2 examines these changes.

I date BPCI to its announcement in 2011, two years before providers began bearing risk. For the post-acute reforms, I use the first rules specifying each payment system: 2000 for home health, 2001 for inpatient rehabilitation facilities, and 2002 for long-term care hospitals. Rules published late in the calendar year are assigned to the following year. The pass-through analysis uses the expiration of separate device reimbursement, while the prescription-to-OTC analysis uses the year of FDA approval. Appendix A reports the announcement and implementation dates.

5.4 Identification and Inference

The identifying assumption is that, absent each reform, the direction of invention would have changed similarly in affected and comparison technologies. This does not require the technologies to have the same initial patent scores. However, differences in scientific opportunities, demand, or other policies could produce changes unrelated to the reforms.

I examine these concerns using event studies that test for differential pre-reform trends and specifications with alternative comparison groups. The preferred matched comparisons

[29]The results are similar when reforms are dated to implementation rather than announcement. Dating BPCI to the start of risk-bearing in 2013 gives an episode-cost estimate of 1.4 percentage points (s.e. 0.4), while dating the post-acute reforms to implementation gives 1.1 (s.e. 0.5). For pass-through expirations and prescription-to-OTC switches, invention responds around the realized change rather than the earlier date at which it became foreseeable.

are selected using pre-reform outcome levels and patenting volume, rather than pre-reform trends, so the event studies provide a separate test of the identifying assumption. For the four reforms that change which costs providers bear, I also allow technology classes to have separate year effects. Appendix Table A.7 reports these checks for each reform.

Standard errors are clustered at the level of the identifying variation. I use a wild cluster bootstrap when there are few treated units, including the FDA Class III-to-II comparison with twelve reclassification events and the competitive-bidding comparison with thirteen treated categories out of 47. Appendix Table A.5 reports the inference procedure for each design.

The reforms also provide separate tests of the mechanism. They differ in timing, affected technologies, and which costs providers bear. Consequently, an alternative explanation must account not only for differential changes in invention, but also for why those changes occur in the particular cost-reducing objectives rewarded by each reform.

6 Results

6.1 Which Reforms Change the Direction of Invention

Table 1 compares the predicted and estimated changes in the direction of invention across reforms. Reforms that make providers bear more of the costs of delivering care shift invention toward reducing those costs. This pattern appears after the 1983 inpatient reform and BPCI, with supporting evidence from three post-acute payment reforms (Table 4). By contrast, FDA reclassification, Medicare competitive bidding, and the medical-device excise tax change the overall returns to invention without changing which costs providers bear. These reforms affect the amount of patenting but produce no statistically significant change in its direction (Table 5 and Figure 3).

The reforms that shift costs to providers also differ in which inventions respond. Episode-based payment increases invention directed toward reducing episode costs, without a corresponding increase in production-cost invention. When separate reimbursement for outpatient devices expires, production-cost invention instead increases while episode-cost invention declines (Figure 4). A general increase in attention to cost reduction would not predict these opposing responses. Appendix Table A.6 reports the effects of each reform on the individual

cost-reducing channels.

The comparison reforms provide a further test. Class III-to-II reclassification increases patenting by approximately 12 patents per product code annually (Rogers, 2026), while competitive bidding substantially reduces patenting (Ji and Rogers, 2024). Neither produces a significant change in the direction of invention. For production-cost invention, the 95 percent confidence intervals rule out increases larger than 0.38 percentage points after Class II-to-I reclassification, 0.55 after competitive bidding, and 0.46 after the device tax, compared with an increase of 1.02 percentage points after pass-through expiry. The Class III-to-II estimate is less precise and cannot rule out an increase of comparable magnitude.[30]

One possible explanation for the competitive-bidding result is that manufacturers respond to lower prices through adjustments other than patented production-cost invention. In my earlier work with Ji and Rogers (2024), we find that competitive bidding reduced domestic innovation and increased reliance on foreign production. Firms may therefore respond to price reductions by changing where they manufacture, reducing margins, or exiting, rather than developing new technologies that lower production costs. This distinction is consistent with the framework, in which changes in the overall returns to invention need not change its direction.

6.2 The 1983 Prospective Payment System

I begin with Medicare's 1983 Inpatient Prospective Payment System (PPS), which replaced cost-based reimbursement with a fixed payment per admission based on diagnosis-related groups (DRGs). Hospitals consequently became responsible for the marginal costs of treating patients during a stay. However, the financial pressure to reduce costs increased over time. Medicare inpatient operating margins were initially above 13 percent, but fell below zero by 1990–1991 as payment updates lagged behind rising costs.[31]

[30]The upper confidence bound for Class III-to-II reclassification is 2.70 percentage points. On the episode-cost score, Class II-to-I reclassification has an estimate of -0.51 percentage points (s.e. 0.45), while competitive bidding has an estimate of $+0.47$ (s.e. 0.36). Both are statistically insignificant (Appendix Figure A.1). Comparing the episode-cost estimates with BPCI, the differences are statistically significant for Class II-to-I reclassification ($z = 3.22$) and the device tax ($z = 3.76$), but not competitive bidding ($z = 1.83$) or Class III-to-II reclassification ($z = 1.03$). On the production-cost margin, the estimates for Class II-to-I reclassification ($z = 2.63$), the device tax ($z = 2.54$), and competitive bidding ($z = 2.53$) differ significantly from the pass-through estimate. These comparisons treat estimates from different designs as independent.

[31]Average Medicare PPS operating margins were 13.5% in FY1984 and 13.7% in FY1985, falling to -2.4% in 1991, when 61% of hospitals had negative margins. Margins subsequently recovered to 16.9% in 1997

Hospitals initially offset some of Medicare's payment reductions by charging higher prices to private insurers. As managed care expanded in the early 1990s, hospitals became less able to do so, increasing the pressure to reduce their own costs (Cutler, 1998). Consistent with this change, Medicare inpatient costs per discharge, which had grown by 8.2 percent in 1990, declined in each year from 1994 to 1996 (Medicare Payment Advisory Commission, 2002). I therefore estimate the invention response separately for 1983–1990 and 1991–1995. The latter period captures stronger financial pressure under the payment system introduced in 1983, rather than a separate reform.

I compare patents in 27 CPC main groups of technologies used mainly in inpatient care with patents in 30 groups used mainly outside it, using filings from 1977 to 1995 (Section 4). The outcome is the average score for inventions that reduce complications or shorten hospital stays, the two channels rewarded by fixed payments per admission. I estimate separate treatment effects for 1983–1990 and 1991–1995, relative to 1977–1982, with technology and filing-year fixed effects and standard errors clustered by technology field.

Figure 2, panel (a), shows that inpatient technologies shift toward reducing episode costs after 1983, with a larger response after 1990. The episode-cost score increases by 1.0 percentage point during 1983–1990, although the estimate is not statistically significant, and by 3.3 percentage points during 1991–1995 (Table 4). The difference between the periods is 2.3 percentage points (s.e. 1.0).[32] The pre-reform coefficients are jointly statistically insignificant ($p = 0.15$). The difference begins to emerge in 1985, remains between 0.9 and 1.7 percentage points through 1990, and reaches approximately 4 percentage points by 1995. The 1991–1995 increase equals 1.5 standard deviations of changes in the comparison technologies over the same period (Section 7.4).

The quality-improvement score falls by 1.7 percentage points over 1991–1995, although this estimate is not statistically significant (s.e. 1.7). Thus, the increase in episode-cost invention is not accompanied by a corresponding increase in quality-improving invention.

One concern is that the response coincides with the spread of minimally invasive surgery

(MedPAC, Report to the Congress, March 2002, Table B-4; Sheingold and Richter 1992; Medicare Payment Advisory Commission 1999).

[32]The estimates are +0.0103 (s.e. 0.0072, $p = 0.16$) for 1983–1990 and +0.0331 (s.e. 0.0129, $p = 0.013$) for 1991–1995. The difference has $p = 0.022$ and remains statistically significant when I divide the periods at any year from 1987 to 1994. Using a continuous measure of inpatient exposure gives an increase of 1.7 percentage points per standard deviation of exposure over 1991–1995 (s.e. 0.6; Appendix C).

(MIS), including laparoscopic gallbladder removal, which expanded rapidly after 1989 (Steiner et al., 1994). These technologies shorten hospital stays and may themselves have become more valuable under prospective payment. However, the increase in episode-cost invention is not confined to fields where minimally invasive surgery was feasible. I measure each technology field's suitability for MIS using a language-model assessment that does not mention Medicare or the payment reform. Excluding the 13 fields rated most suitable for MIS leaves estimated increases of 1.4 percentage points during 1983–1990 and 3.9 percentage points during 1991–1995, with no significant pre-reform differences.[33]

A second concern is the introduction of capital prospective payment in 1992, which gradually incorporated hospitals' capital costs into DRG payments. However, this reform would more directly increase incentives to reduce equipment purchase costs, rather than complications or length of stay. The shift toward episode-cost invention also began well before 1992 and persists outside fields suitable for MIS. Dividing the sample at 1988, before the capital rules were enacted, still produces a significantly larger response in the later period.[34]

6.3 Bundled Payment and Other Payment Reforms

I next examine whether similar changes occur under other payment reforms, and whether the particular costs providers bear determine which inventions respond.

BPCI provides a separate test of episode-based payment nearly three decades after the 1983 reform. Announced in 2011, the program made participating providers responsible for spending across 48 clinical episodes, including joint replacement. I compare patents linked to covered procedures with patents linked to similar procedures outside the program. After the

[33] I average five ratings from gpt-5-mini for each CPC main group. The prompt asks whether procedures in the field can be performed through scopes, catheters, or small incisions rather than open operations, without mentioning payment, Medicare, hospitals, or length of stay. The mean pairwise correlation across ratings is 0.93. Excluding fields with scores of at least 0.7 removes 11 inpatient and two comparison fields. The estimates are +0.014 (s.e. 0.006) and +0.039 (s.e. 0.012), with a joint pre-reform test of $p = 0.30$. Allowing MIS suitability to have separate effects in each period gives an inpatient estimate of +0.026 (s.e. 0.013, $p = 0.058$) for 1991–1995, while the coefficient on MIS suitability is +0.005 (s.e. 0.011, $p = 0.62$). Because MIS may itself respond to payment incentives, controlling for its feasibility may absorb part of the reform's effect.

[34] The later estimate is 2.65 percentage points, compared with 0.84 in the earlier period ($p = 0.03$ for the difference). Appendix A examines the timing of capital PPS and reports a structural-break test for durable-equipment technologies. The test finds no break in 1992, although it is imprecise given the gradual phase-in.

announcement, the episode-cost score of exposed patents increases by 1.6 percentage points relative to comparison patents (Table 4).[35] Figure 2, panel (b), shows no differential pre-reform trend. The response begins after the 2011 announcement and increases after providers began bearing financial risk in 2013, from approximately 0.9 percentage points during 2011–2013 to 1.6 percentage points from 2014 onward. Because BPCI participation was voluntary, the estimate averages over technologies used by both participating and nonparticipating providers.[36]

The three post-acute reforms provide further evidence. Medicare introduced prospective payment for home health agencies in 2000, inpatient rehabilitation facilities in 2001, and long-term care hospitals in 2002. Each reform made providers responsible for costs within its payment unit. Comparing affected device fields with matched acute-care technologies, I estimate an increase of 1.1 percentage points in the episode-cost score (s.e. 0.5; Appendix Figure A.2).[37] The response begins around implementation and grows to approximately 1.9 percentage points six years later. Exposure is measured at the technology-field level, since patent data do not identify where each device is used. Some devices within treated fields are therefore used outside post-acute care, which may attenuate the estimate.

The episode-cost estimates are similar across reforms despite differences in the affected technologies and payment systems. Measured relative to changes in unaffected technologies over the same periods, the estimated increases are 1.5 standard deviations for the 1983 reform during 1991–1995, 0.9 for BPCI, and 1.0 for the post-acute reforms (Figure 3). The payment units also differ substantially in size, from inpatient admissions costing several thousand dollars to post-acute stays of approximately $16,000 and BPCI joint-replacement episodes of approximately $27,000–31,000 (Appendix Table A.8).[38] The similarity in the

[35]The estimate is +0.0158 (s.e. 0.0046, $p = 0.001$). The pre-reform coefficients are jointly insignificant ($p = 0.97$). Using the share of each patent's procedure links covered by BPCI as a continuous measure of exposure gives an increase of 1.0 percentage point at the average exposure of treated patents (Appendix C).

[36]The Comprehensive Care for Joint Replacement program, announced in 2015 and implemented in selected metropolitan areas in 2016, covered episodes already included in BPCI. I therefore do not estimate its effect separately.

[37]The estimate is +0.0111 (s.e. 0.0048, $p = 0.023$), with pre-reform coefficients jointly insignificant ($p = 0.996$). Using all three episode-cost channels for every setting gives +0.0116 (s.e. 0.0046). Skilled nursing facilities are excluded because their per-day payment does not directly reward shorter stays. Inpatient psychiatric facilities are excluded because they are not a post-acute setting.

[38]The post-acute estimate is $16,649, based on MedPAC data. BPCI joint-replacement episode spending ranges from $27,265 for intervention episodes to $30,551 for baseline episodes (Dummit et al., 2016). The 1983 DRG payment magnitude is based on the CMS FY1984 rule.

invention responses is consistent with the importance of which costs providers bear, rather than the dollar size of the payment alone.

I next examine the expiration of outpatient pass-through payments, which changes providers' exposure to device purchase prices rather than the broader costs of care. Medicare provides separate reimbursement for qualifying new outpatient devices for two to three years. Once this payment expires, the device cost is included in the fixed payment for the procedure, giving hospitals an incentive to purchase less expensive devices. Because the expiration date is determined when separate reimbursement begins, manufacturers may respond before the payment actually expires. Alternatively, invention may respond primarily when hospitals begin bearing the device's cost.

Using device categories that received pass-through payments between 2000 and 2018, I compare patent scores before, during, and after separate reimbursement. Appendix Figure A.3 shows no significant change in production-cost scores when separate reimbursement begins. After it expires, the production-cost score rises by 1.0 percentage point relative to the years of separate reimbursement, while the episode-cost score falls by 0.9 percentage points.[39]

The production-cost response remains positive when I compare patents filed by the same firm in the same year (+0.6 percentage points, $p = 0.04$), suggesting that it is not driven entirely by changes in which firms patent. It also remains when I control for the age of each device category. The decline in episode-cost invention is consistent with a shift away from that objective, although the within-firm estimate is not statistically significant.

These reforms provide a more specific test of the mechanism than a general shift toward cost reduction. When providers bear more of the costs of delivering care, invention shifts toward reducing episode costs. When they instead bear device purchase prices, invention

[39] I stack device categories by pass-through start year (eleven cohorts), comparing their patents with those linked to 6,888 billing codes that never received pass-through payments. The estimates include code and year fixed effects within each cohort, weight code-year observations by their patent counts, and cluster standard errors by code. For production cost, the estimate is -0.0005 ($p = 0.88$) during separate reimbursement relative to the preceding period, and $+0.0102$ ($p = 0.003$) after expiration relative to separate reimbursement. The corresponding episode-cost estimate is -0.0091 ($p = 0.015$). The pre-reform coefficients are jointly insignificant ($p = 0.93$ and 0.78). Controlling for time since pass-through began gives estimates of $+0.0112$ ($p = 0.049$) and -0.0183 ($p = 0.007$). Approximately two-thirds of treated patents are associated with categories whose payments expired at the end of 2002. Excluding these categories reduces the estimates to $+0.0070$ ($p = 0.28$) and -0.0046 ($p = 0.59$). Under three alternative classifiers, the production-cost estimate remains positive ($+0.0010$ to $+0.0055$), but is statistically significant only under the baseline classifier.

shifts toward reducing production costs. Thus, the particular costs that payment reforms shift onto providers are associated with the particular types of cost-reducing inventions that follow.

I exclude labor saving from the primary episode-cost measure because I cannot validate it against corresponding changes in care delivery. However, labor-saving invention also increases after the 1983 reform, by 3.0 percentage points during 1991–1995, and under BPCI, by 2.2 percentage points. I find no significant response in the post-acute design or within firms under BPCI (Appendix Table A.6).

Finally, I examine whether these changes occur within firms or reflect changes in which firms patent. Figure 5 compares patents of the same firm filed in the same year, holding constant firm-wide changes in invention. Under BPCI, exposed patents shift toward episode-cost reduction by 1.6 percentage points relative to the firm's other patents (s.e. 0.35). The response grows after 2011 and averages approximately 2 percentage points during 2014–2020. On the same sample, the estimate without firm-by-year fixed effects is 2.1 percentage points, so most of the response among incumbent firms remains when I compare patents filed by the same firm in the same year.[40]

The within-firm estimate does not fully separate changes in firms' own invention from changes in which firms account for exposed patenting. I therefore decompose the change in exposed patenting, relative to unexposed patenting, into changes within firms and changes in the distribution of patenting across firms. Approximately 32 percent of the increase comes from changes within surviving firms, 38 percent from surviving firms shifting their patenting toward exposed technologies, and 29 percent from firms entering the exposed segment from other device fields. New patentees account for little of the change.[41]

The results are also unlikely to reflect changes in patent language alone. The estimates

[40]The firm-by-year estimate is +0.0164 (s.e. 0.0035), with technology-class and firm-by-year fixed effects, standard errors clustered by firm and technology class, and jointly insignificant pre-reform coefficients ($p = 0.44$). Controlling for pre-reform differences between each firm's exposed and unexposed patents gives +0.0132 (s.e. 0.0047). I estimate within-firm effects for BPCI and pass-through expiry because many exposed patents are filed by firms with comparison patents in the same year. This overlap is much smaller for the 1983 and post-acute reforms. Across firms' entire portfolios, the BPCI response is too diluted to detect, since exposed patents account for approximately 1% of a typical firm's patenting (estimate +0.0003, $p = 0.81$).

[41]The decomposition follows Griliches and Regev (1995), weighting each firm by its average share of exposed patents in the two periods. Within surviving firms: 32% (90% confidence interval 2–55%); reallocation among surviving firms: 38% (−22–89%); firms entering from other device fields: 29% (10–57%); new patentees: 2%; patents without a company assignee: 7%; and exit: −8%. Using base-period shares gives a within-firm contribution of 36%. Confidence intervals are bootstrapped over primary procedure codes.

remain similar when I compare patents filed by the same firm or drafted by the same law firm in the same year.[42] Excluding firms that begin patenting after each reform leaves at least three-quarters of the estimated effects for the 1983, BPCI, and post-acute reforms. The BPCI and pass-through estimates also persist within finer technology groups, at 1.3 and 0.7 percentage points, respectively (Appendix Table A.9). Together, these results indicate that existing firms redirect invention within affected technologies, while changes in the distribution of patenting across firms also contribute.

7 Welfare and Incidence

The patent measures identify the objectives of invention, rather than the savings those inventions ultimately produce. I therefore examine whether the redirection is associated with changes in the amount of invention, manufacturers' costs and device prices, and the costs and quality of care. These outcomes also provide evidence on who benefits from different types of cost reduction, although they do not identify the welfare effects of individual patents.

7.1 The Amount of Invention

The shift toward cost-reducing invention was not accompanied by a decline in patenting. In the 27 inpatient technology fields, patenting increased relative to the 30 comparison fields after 1983, by 20 percent during 1983–1990 (95 percent confidence interval: −1 to +44 percent) and 79 percent during 1991–1995 (+24 to +158 percent). The increase follows the same timing as the shift in the direction of invention. There is no significant pre-reform trend, although the joint test is marginally significant because of an increase in 1979.

The shift in direction is also present among fields where patenting grew less. Among inpatient fields with below-median patent growth, the 1991–1995 increase in the episode-cost score is 2.6 percentage points (s.e. 1.0), compared with 3.5 percentage points (s.e. 1.5) among faster-growing fields. The difference is not statistically significant.[43] Patenting is also

[42] Among firm-assigned patents, the estimate is 1.88 percentage points without firm effects and 1.50 with firm-by-year effects, compared with 1.58 in the baseline. Among the 87 percent of patents linked to law firms in PatentsView, the estimates are 1.53 without and 1.75 with law-firm-by-year effects.

[43] I estimate Poisson regressions of patent counts by technology field and filing year from 1977 to 1995, with field and year fixed effects and standard errors clustered by field. The estimates are +0.179 (s.e. 0.094) for 1983–1990 and +0.582 (s.e. 0.186) for 1991–1995. Matching each inpatient field to its three nearest

approximately unchanged following the post-acute reforms, with confidence intervals ruling out a decline larger than 22 percent. Thus, the observed redirection does not appear to come at the expense of fewer patented inventions in affected technologies.

7.2 Production Costs and Device Prices

As shown in Section 4, firms that redirect their patenting toward reducing production costs also have lower manufacturing-cost ratios around the time of patent filing. However, lower production costs need not translate into lower device prices. Whether manufacturers retain the savings or pass them on to buyers depends on how prices respond. I therefore examine manufacturers' prices, changes in hospitals' purchasing decisions, and device prices around patent expiration.

The lower cost ratio is not accompanied by a corresponding decline in manufacturers' own device prices. A manufacturer's production-cost score has no statistically significant association with changes in its posted prices (-1.5 percent per standard-deviation-year, $p = 0.60$). However, the prices hospitals pay can fall through changes in which products they purchase. In device categories with more production-cost invention, the quality-adjusted average price paid by hospitals declines approximately 3 to 4 percent faster per standard deviation of the production-cost score each year, depending on the quality controls (Appendix Figure A.4). This difference reflects changes in the composition of purchases rather than price reductions for existing products. When I weight products equally, the difference is only -0.3 percent ($p = 0.65$). Most of the change comes from hospitals purchasing cheaper products already on the market. Changes in hospital budgets or contracting may also contribute to this relationship (Appendix D).

One explanation is that manufacturers retain part of the production-cost savings while their patents remain in force. If a device's price reflects its value to the hospital rather than its manufacturing cost, an invention that lowers production costs need not reduce its price. The manufacturer can instead retain the savings as a higher margin. The firm-level results are consistent with this mechanism, although they do not establish how much of the savings

comparison fields on patenting during 1977–1982 gives +0.180 (s.e. 0.085) and +0.599 (s.e. 0.120). The pre-reform slope is -0.01 per year ($p = 0.74$); the joint test has $p = 0.08$, reflecting a rise in 1979. I measure patent growth as the change in log mean annual patent counts from 1977–1982 to 1991–1995. The difference in direction between fields with above- and below-median growth has $p = 0.55$.

manufacturers retain.

I examine whether device prices fall when production-cost patents expire. Patent terms are generally determined by their filing dates, approximately 20 years before expiration, making the timing less likely to respond to contemporaneous changes in device prices. Using the patent-to-product-code crosswalk, I identify 24 device categories with large expirations of production-cost patents and compare their prices with matched categories without large nearby expirations (Appendix Figure A.5). Prices show no significant differential trend before expiration and fall by approximately 7 percent in the expiration year, although the estimate is imprecise.[44] Expirations of quality-improving patents are associated with a smaller decline, but the two types of expiration overlap and their estimates are not statistically distinguishable (Appendix Figure A.6).

The decline in average prices again reflects changes in purchases rather than lower prices for continuing products. In the expiration year, hospitals shift toward cheaper products already available, while incumbent manufacturers introduce cheaper models. New manufacturers account for little of the change. Quantities increase by approximately 14 percent, so hospitals purchase more devices at a lower average price. The decomposition is less precise in later years (Appendix Figure A.7). These patterns suggest that production-cost savings may reach buyers after patent expiration, although the estimates do not establish that the expiration itself caused the price decline.

I also examine whether prices fall when hospitals become responsible for device purchase costs. Medicare initially reimburses some new devices separately through outpatient pass-through payments or inpatient new-technology add-on payments. Once these payments expire, hospitals must cover the device's price from their fixed reimbursement. If greater cost exposure strengthens hospitals' bargaining incentives, their purchase prices may decline.

Using ECRI transaction data, I compare five devices whose separate payments expired with similar products purchased by the same hospitals. After accounting for each device's price trend relative to its comparison products, the estimated price change is -0.7 percent, with a 95 percent confidence interval of -2.0 to $+0.6$ percent. Across the five devices, the

[44]The expiration-year estimate is -0.074 log points (s.e. 0.059), and the average over the expiration year and following three years is -0.103 (s.e. 0.077). The pre-expiration coefficients are jointly insignificant ($p = 0.94$). Appendix D reports alternative comparison groups and discusses variation between nominal filing-plus-20-year terms and actual expiration dates.

equally weighted change is +0.7 percent, although individual changes range from −8 to +15 percent.[45]

I therefore find little evidence that greater exposure to device prices immediately lowers the prices hospitals pay, although the small number of devices and possible unobserved rebates limit this conclusion. More generally, the absence of substantial price changes among continuing products is consistent with manufacturers retaining some production-cost savings as margin. I cannot determine how much of this pattern reflects market power, since interactions between production-cost invention and manufacturer concentration are imprecise and differ across concentration measures (Appendix D).[46]

7.3 Costs of Delivering Care

I next examine whether the shift toward reducing episode costs is associated with changes in care delivery. The 1983 reform provides the longest period over which I can observe these outcomes. Hospital stays declined substantially after the reform, including among privately insured patients who were not directly subject to Medicare's new payment system. This pattern is consistent with hospitals adopting technologies and practices that reduce the costs of care for all patients, rather than changing discharge decisions only for Medicare patients.

Using the National Hospital Discharge Survey (NHDS), I compare changes in length of stay within procedures across age groups and payers (Appendix Figure A.8, panel (a)). Stays decline by similar amounts among patients younger and older than 65. Comparing

[45]The ECRI data cover hospital purchases from 2012 to 2019. The devices are drug-coated angioplasty balloons, the Zilver PTX stent, the INTUITY and Perceval aortic valves (treated as one device), a powered bone-marrow biopsy system, and a lung biopsy plug. Comparison products are in the same device class, were never paid separately, and were purchased by the same hospitals. The regressions include product-by-hospital and quarter effects and remove each device's linear price trend relative to its class. Standard errors clustered by hospital give a 95% confidence interval of $[-2.0\%, +0.6\%]$ for the pooled estimate, but do not account for price shocks common to a device. The between-device confidence interval for the equal-weighted change is $[-9\%, +11\%]$. Without removing relative price trends, the estimate is +2.0% (95% CI $[+0.9\%, +3.1\%]$; wild cluster bootstrap by device, $p = 0.11$). Four additional devices show little change in catalog prices. The angioplasty balloon and stent samples end before the 2018 paclitaxel safety signal. Invoice prices do not capture rebates paid afterward.

[46]The interaction changes sign when concentration is measured using the four-firm concentration ratio rather than the Herfindahl index. In the ECRI data, a catalog item's price at a hospital is unchanged from one year to the next approximately half the time, with an average price spell of 2.4 years. However, 41% of prices change by more than 1 percent in a year, by a median of 5.6%. Appendix D provides additional details.

patients of the same age, I also find no differential change in length of stay between Medicare and privately insured patients. Although a comparison of all patients younger than 65 with those 65 and older suggests a 2.5 percent larger decline among older patients, this difference disappears when I hold age fixed. I therefore do not attribute the decline in hospital stays specifically to Medicare patients being discharged sooner.

Minimally invasive surgery provides one example of a technology that shortened hospital stays during this period. Laparoscopic and endoscopic procedures expanded rapidly following the reform, replacing more invasive operations and reducing recovery times. These technologies are consistent with the direction of invention identified by the patent measure, although the data do not establish how much of the decline in length of stay was caused by the inventions induced by prospective payment. The patenting response also extends beyond fields suitable for minimally invasive surgery (Section 6.2). Consistent with a broader change in hospital technology, Acemoglu and Finkelstein (2008) find that hospitals with greater Medicare exposure adopted new technologies faster and experienced larger declines in length of stay.

Medicare patients did experience a separate change in where they received care. Comparing Medicare and privately insured patients aged 55–74 with the same procedure, transfers to other facilities increased by 0.9 percentage points after 1983, while length of stay and in-hospital mortality showed no significant differential changes (Appendix Figure A.8).[47] Because Medicare reimburses post-acute care separately, transferring patients to another facility can move costs outside the hospital's fixed payment. The increase in transfers is therefore consistent with hospitals responding to the costs they bear.

The mortality evidence requires a qualification. NHDS records deaths during hospitalization but not after discharge. Consequently, unchanged in-hospital mortality does not rule

[47] I use NHDS discharges from 1979–1990, comparing patients whose expected principal payer is Medicare with those covered by commercial insurance other than Blue Cross. The regressions include procedure-by-payer and procedure-by-year fixed effects, weight observations by discharges, and cluster standard errors by procedure. The Medicare × post-1983 estimates are +0.87 percentage points for transfers ($p < 0.001$), −0.000 for log length of stay (95% confidence interval ±0.024), and −0.003 percentage points for in-hospital mortality ($p = 0.99$). The transfer effect is also positive among patients aged 60–64 (+1.17, $p = 0.035$). Because the 1979 transfer coefficient is unusually high, the pre-reform coefficients are not jointly insignificant. Excluding 1979 gives +1.14 percentage points (pre-trend $p = 0.13$), while allowing a separate linear trend for Medicare patients gives +0.71 ($p = 0.035$). Within five-year age groups, the transfer estimate is +0.81 ($p = 0.086$). Including Blue Cross in the comparison group gives +0.86 ($p < 0.001$), and starting in 1981, after NHDS changed its discharge-status codes, gives +1.29 ($p < 0.001$).

out patients being discharged earlier or transferred before death. The data establish that stays shortened across payers and that transfers increased among Medicare patients, but cannot determine whether the reform changed overall mortality.

For modern bundled payments, I use published evaluations to examine spending and quality over the full episode of care. The Comprehensive Care for Joint Replacement (CJR) program covers procedures also included in BPCI, although it is a separate payment reform. In a two-year difference-in-differences evaluation, Barnett et al. (2019) find that episode spending declined by $812, or 3.1 percent, primarily through reduced institutional post-acute care. Complications did not increase, and readmissions were slightly lower. However, this estimate excludes reconciliation payments to participating hospitals.

A randomized evaluation of CJR by Finkelstein et al. (2018) finds a 2.9-percentage-point reduction in discharges to institutional post-acute care, with no statistically significant deterioration in complications, readmissions, or a composite quality measure that includes mortality. Total episode spending declines by $453, an estimate that is not statistically significant. After reconciliation payments, the estimated effect on Medicare spending is instead +$234, also statistically insignificant.[48]

These evaluations show that bundled payment can reduce the use of costly post-acute settings without a detectable decline in measured quality. However, the reduction in spending before reconciliation payments need not translate into equivalent Medicare savings, since participating hospitals may receive some of the savings through incentive payments. The evidence therefore distinguishes lower resource use from savings retained by the payer.

Taken together, the evidence is consistent with the two types of cost-reducing invention producing savings through different channels. Episode-cost invention is associated with technologies that shorten hospital stays or allow care to move to less costly settings. Production-cost invention is associated with lower manufacturing-cost ratios and higher margins, with more limited evidence of lower prices for continuing products. These patterns help explain the potential economic importance of the redirection, but I cannot attribute the observed changes in care delivery or device prices to individual patents.

[48]In Barnett et al. (2019), the complication estimate has $p = 0.67$. In Finkelstein et al. (2018), the complication estimate has $p = 0.99$ and the composite quality estimate has $p = 0.10$. The latter study reports the −$453 spending estimate before reconciliation payments and +$234 afterward; neither differs significantly from zero.

7.4 How Far Payment Design Reaches

The model predicts that increasing the share of costs providers bear shifts invention toward reducing those costs (Appendix B). I examine the magnitude of these responses relative to typical changes in invention and the share of device technologies that remain outside prospective payment.

First, I compare each reform's estimated effect with the standard deviation of changes in unaffected technologies over the same period. This benchmark does not depend on the level of the classifier score, which is not calibrated to the prevalence of each invention objective. The 1983 reform increases the episode-cost score by 1.5 standard deviations of changes in unaffected fields over 1991–1995. The corresponding effects are 0.9 standard deviations for BPCI and 1.0 for the post-acute reforms. The production-cost response to pass-through expiration is 0.8 standard deviations.[49]

Second, I examine the share of patented device technologies that remain outside prospective payment. Using the patent-to-HCPCS crosswalk, I classify device patents according to their associated payment settings. Approximately three-quarters are linked to settings paid prospectively under the inpatient and outpatient payment systems. The remaining 26 percent are linked to the durable medical equipment and orthotics fee schedule or outpatient device pass-through payments. This share rises to 44 percent when supplies are included.[50]

These shares suggest that payment design could affect invention beyond the technologies already exposed to prospective payment. Under the narrow definition, the number of patents associated with technologies outside prospective payment is approximately one-third as large

[49]For each design, I apply its before-and-after comparison to untreated units over feasible start years, net of unit and year means. I calculate the standard deviation of these changes after accounting for sampling noise, weighting by patent counts. The benchmark for the 1983 reform is 2.2 percentage points, calculated from 22 device fields used mainly outside inpatient care with at least 30 patents in each period. These include dental technologies, which were excluded from the estimation comparison group but were unaffected by the reform. The 3.3-percentage-point effect is 1.52 standard deviations, while the 1983–1990 effect is 0.67. For BPCI, the benchmark is 1.8 percentage points across 95 unaffected technology classes, giving 0.89 standard deviations. The post-acute estimate is 0.97 standard deviations, using 29 matched comparison fields with sufficient patents and the same mix of outcomes as the treated fields. For pass-through expiration, the benchmark uses 253 untreated codes and gives 0.81 standard deviations. Alternative comparison groups and windows produce ranges of 1.43–1.69 for 1983, 0.85–1.11 for BPCI, 0.76–0.97 for the post-acute reforms, and 0.67–0.81 for pass-through expiration. Under PaECTER, the corresponding values are 1.49, 0.71, 0.65, and 0.48 (Appendix Table A.10).

[50]Appendix D describes the classification based on cosine-weighted HCPCS-II links (26.0% under the narrow definition and 44.0% under the broader definition), the refinement using CPT place-of-service information, and why I do not construct an equivalent measure based on spending.

as the number already associated with prospectively paid settings. If extending cost exposure produces similar responses across technologies, the potential additional redirection could therefore be substantial. However, this comparison measures the distribution of patents rather than spending or the costs providers would bear under a new payment system.[51]

8 A Demand-Side Extension

I next examine whether payment incentives also shape invention when patients, rather than providers, choose and pay for new technologies. Medical devices provide little variation for this question. Of the 1,052 FDA device types in the patent panel, patients choose only 51, nearly all of which are purchased out of pocket (47). None is typically covered by insurance. For insured devices, physicians, hospitals, or suppliers generally make the purchasing decisions (Appendix A).

Prescription-to-over-the-counter (OTC) drug switches provide a setting in which patients become both the payer and the decision-maker. The switch itself reduces the cost of obtaining a drug by eliminating the need for a prescription. The model instead predicts a change in subsequent invention: when patients choose and pay for drugs themselves, inventions that reduce their remaining costs of use become more valuable. These include less frequent dosing, easier self-administration, and reduced monitoring requirements.[52]

For this analysis, I construct a separate measure of drug inventions that reduce patients' costs of use. Patents associated with new dosage forms or formulations, which can make drugs easier to take, score 1.4 standard deviations higher on this measure than patents associated with new molecular entities (Figure 6, panel (a)).

[51] This comparison is not an estimate of the welfare gains from extending prospective payment. The evidence for responses beyond inpatient care comes from the BPCI, post-acute, and pass-through reforms. These reforms affect different technologies and costs, yet each produces a shift toward the type of cost reduction rewarded by its payment rule. I do not translate these shifts into dollar benefits because the patent measures do not identify the savings generated by individual inventions. The evidence instead indicates that payment design can change which types of technologies are developed, with implications for both the costs of care and who benefits from cost reduction.

[52] Medicare Part D provides a related comparison. The reform shifted outpatient drug costs from patients to private plans without transferring prescribing decisions to patients. I find no significant change in invention directed toward reducing patients' costs of use. The estimates are +0.3 percentage points (s.e. 1.2) when comparing self-administered with physician-administered drugs, and −0.2 percentage points (s.e. 0.9) per standard deviation of a drug's 2001–2003 Medicare spending share (Duggan and Scott Morton, 2010). Both designs have flat pre-reform trends, although the estimates are imprecise.

Across 24 OTC switches, the patient-use-cost score increases by 3.6 percentage points (s.e. 1.4) after FDA approval, or 1.2 standard deviations of changes in drugs that never switched. Figure 6, panel (b), shows no differential trend before the switch, with the increase appearing two to five years afterward. I date each switch to FDA approval rather than the preceding application or advisory review, since approval is not guaranteed. The response is also concentrated after approval rather than these earlier signals.[53]

Because patients also bear the drug's purchase price after an OTC switch, the model predicts an increase in production-cost invention. I find no significant change (+0.5 percentage points, s.e. 1.8), but the pre-reform coefficients are not flat (joint $p = 0.013$), so this design does not provide reliable evidence on that margin.

The increase in patient-use-cost invention is consistent with the mechanism extending to patients: when patients become responsible for choosing and paying for drugs, invention shifts toward reducing their costs of use. However, OTC switches also change how drugs are purchased and used, so I cannot separate the effects of payment incentives from these other changes.

9 Discussion and Policy Implications

This paper examines how payment incentives shape the direction of medical invention. I find that reforms that make providers bear more of the costs of care shift invention toward reducing those costs. After Medicare introduced prospective payment in 1983, invention in inpatient technologies shifted toward reducing episode costs by 3.3 percentage points over 1991–1995. Similar shifts followed bundled payment nearly three decades later (1.6 percentage points) and three post-acute payment reforms (1.1 percentage points). By contrast, when separate reimbursement for outpatient devices expires and providers become responsible for device purchase prices, invention shifts toward reducing production costs (1.0 percentage

[53]Each treated molecule is the ingredient form that switched to OTC status. Patents linked to multiple molecules are assigned to the molecule with the highest text similarity. The pre-reform coefficients are jointly insignificant ($p > 0.99$). Dating switches three or five years earlier gives smaller, statistically insignificant estimates of +2.3 and +0.8 percentage points, respectively. Using earlier announcement dates and rematching comparison drugs gives an estimate of +0.5 percentage points. OTC switches sometimes require new strengths, dosage forms, or labels, so some inventions may reflect regulatory requirements rather than the change in payment incentives. However, formulations required for approval must exist before the switch, whereas the increase in patenting occurs afterward. Appendix C describes the design.

point), while invention directed toward reducing episode costs declines. Reforms that change the overall returns to innovation without changing providers' cost exposure affect patenting but produce no comparable change in these measures of its direction.

The results suggest that payment design affects not only whether inventors pursue cost reduction, but which costs they seek to reduce. This pattern appears across reforms spanning more than three decades and different technologies. The response occurs partly within firms, as manufacturers redirect invention in affected technologies, and partly through changes in which firms patent. The response to the 1983 reform also extends beyond technologies suitable for minimally invasive surgery, suggesting that it is not driven solely by new opportunities in those fields. Together, these findings support the mechanism in which payment incentives change the value of different improvements to technology adopters, thereby changing the relative returns to developing them.

These findings have implications for policies aimed at reducing health care spending. One approach is to subsidize inventions that lower costs through research grants, prizes, or tax credits. Such subsidies could increase cost-reducing invention, but whether the savings reach buyers depends on how device prices respond. I find that firms with more production-cost invention have lower manufacturing-cost ratios and higher gross margins, without corresponding reductions in the prices of continuing products. Device prices decline after production-cost patents expire, although that estimate is imprecise. These results suggest that manufacturers may retain some production-cost savings while their patents remain in force. Directly reducing device prices could instead transfer savings to buyers, but may also reduce the returns to developing new technologies, as illustrated by the decline in innovation following Medicare competitive bidding (Ji and Rogers, 2024).

Prospective and bundled payments operate through a different incentive. By making providers responsible for the costs of delivering care, these payment systems increase the value of technologies that reduce those costs. Rather than rewarding manufacturers directly for cost-reducing invention, they change what providers are willing to pay for. The evidence suggests that this can redirect invention toward technologies that shorten hospital stays, prevent complications, or allow treatment in less costly settings. However, the benefits depend on whether these changes reduce the overall costs of care rather than simply shifting costs across providers or payers, and whether quality is maintained.

Several limitations qualify these conclusions. The patent measures identify the objectives of invention, not the savings generated by individual patents. Although the production-cost measure is associated with lower manufacturing costs and the episode-cost measures track changes in care delivery, I cannot attribute those outcomes to the inventions induced by payment reform. The post-acute designs also measure exposure using broad technology fields, while the classifier introduces measurement error that the robustness tests cannot fully rule out. Moreover, the evidence on who receives production-cost savings is limited by incomplete transaction-price data and imprecise patent-expiration estimates. The broader applicability of the mechanism also remains uncertain: the OTC-switch analysis provides suggestive evidence that it extends to patients, but those switches change more than who chooses and pays for the drug.

The central implication is that payment systems can influence the direction of technological progress by changing which improvements adopters value. When providers bear the costs of delivering care, invention shifts toward reducing those costs. When they bear device purchase prices, invention shifts toward reducing production costs. Thus, the particular costs that payment systems assign to technology adopters help determine which kinds of medical inventions are developed.

References

Acemoglu, Daron, "Directed Technical Change," *Review of Economic Studies*, 2002, *69* (4), 781–809.

_ **and Amy Finkelstein**, "Input and Technology Choices in Regulated Industries: Evidence from the Health Care Sector," *Journal of Political Economy*, 2008, *116* (5), 837–880.

_ **and Joshua Linn**, "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry," *Quarterly Journal of Economics*, 2004, *119* (3), 1049–1090.

Aghion, Philippe, Antoine Dechezleprêtre, David Hémous, Ralf Martin, and John Van Reenen, "Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry," *Journal of Political Economy*, 2016, *124* (1), 1–51.

Azoulay, Pierre, Joshua S. Graff Zivin, Danielle Li, and Bhaven N. Sampat, "Public R&D Investments and Private-Sector Patenting: Evidence from NIH Funding Rules," *Review of Economic Studies*, 2019, *86* (1), 117–152.

Barnett, Michael L., Andrew Wilcock, J. Michael McWilliams, Arnold M. Epstein, Karen E. Joynt Maddox, E. John Orav, David C. Grabowski, and Ateev Mehrotra, "Two-Year Evaluation of Mandatory Bundled Payments for Joint Replacement," *New England Journal of Medicine*, 2019, *380* (3), 252–262.

Baumgardner, James R., "The Interaction between Forms of Insurance Contract and Types of Technical Change in Medical Care," *RAND Journal of Economics*, 1991, *22* (1), 36–53.

Bena, Jan and Elena Simintzi, "Machines Could Not Compete with Chinese Labor: Evidence from U.S. Firms' Innovation," *Review of Finance*, 2025, *29* (6), 1619–1661.

Budish, Eric, Benjamin N. Roin, and Heidi Williams, "Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials," *American Economic Review*, 2015, *105* (7), 2044–2085.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The Effect of Minimum Wages on Low-Wage Jobs," *Quarterly Journal of Economics*, 2019, *134* (3), 1405–1454.

Chandra, Amitabh and Jonathan Skinner, "Technology Growth and Expenditure Growth in Health Care," *Journal of Economic Literature*, 2012, *50* (3), 645–680.

Clemens, Jeffrey and Parker Rogers, "Demand Shocks, Procurement Policies, and the Nature of Medical Innovation: Evidence from Wartime Prosthetic Device Patents," *Review of Economics and Statistics*, 2026, *108* (1), 75–89.

Cutler, David M., "The Incidence of Adverse Medical Outcomes under Prospective Payment," *Econometrica*, 1995, *63* (1), 29–50.

_ , "Cost Shifting or Cost Cutting?: The Incidence of Reductions in Medicare Payments," *Tax Policy and the Economy*, 1998, *12*, 1–27.

_ **, Allison B. Rosen, and Sandeep Vijan**, "The Value of Medical Spending in the United States, 1960–2000," *New England Journal of Medicine*, 2006, *355* (9), 920–927.

Dafny, Leemore S., "How Do Hospitals Respond to Price Changes?," *American Economic Review*, 2005, *95* (5), 1525–1547.

Dubois, Pierre, Olivier de Mouzon, Fiona Scott Morton, and Paul Seabright, "Market Size and Pharmaceutical Innovation," *RAND Journal of Economics*, 2015, *46* (4), 844–871.

Duggan, Mark and Fiona Scott Morton, "The Effect of Medicare Part D on Pharmaceutical Prices and Utilization," *American Economic Review*, 2010, *100* (1), 590–607.

Dummit, Laura A., Daver Kahvecioglu, Grecia Marrufo, Rahul Rajkumar, Jaclyn Marshall, Eleonora Tan, Matthew J. Press, Shannon Flood, L. Daniel Muldoon, Qian Gu, Andrea Hassol, David M. Bott, Amy Bassano, and Patrick H. Conway, "Association Between Hospital Participation in a Medicare Bundled Payment Initiative and Payments and Quality Outcomes for Lower Extremity Joint Replacement Episodes," *JAMA*, 2016, *316* (12), 1267–1278.

Ellis, Randall P. and Thomas G. McGuire, "Provider Behavior under Prospective Reimbursement: Cost Sharing and Supply," *Journal of Health Economics*, 1986, *5* (2), 129–151.

Finkelstein, Amy, "Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry," *Quarterly Journal of Economics*, 2004, *119* (2), 527–564.

_ , Yunan Ji, Neale Mahoney, and Jonathan Skinner, "Mandatory Medicare Bundled Payment Program for Lower Extremity Joint Replacement and Discharge to Institutional Postacute Care: Interim Analysis of the First Year of a 5-Year Randomized Trial," *JAMA*, 2018, *320* (9), 892–900.

Goddeeris, John H., "Insurance and Incentives for Innovation in Medical Care," *Southern Economic Journal*, 1984, *51* (2), 530–539.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

Griliches, Zvi and Haim Regev, "Firm Productivity in Israeli Industry 1979–1988," *Journal of Econometrics*, 1995, *65* (1), 175–203.

Hanlon, W. Walker, "Necessity Is the Mother of Invention: Input Supplies and Directed Technical Change," *Econometrica*, 2015, *83* (1), 67–100.

Holmstrom, Bengt and Paul Milgrom, "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, & Organization*, 1991, *7* (Special Issue), 24–52.

Ji, Yunan and Parker Rogers, "The Long-Run Impacts of Regulated Price Cuts: Evidence from Medicare," NBER Working Paper 33083, National Bureau of Economic Research 2024.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring Technological Innovation over the Long Run," *American Economic Review: Insights*, 2021, *3* (3), 303–320.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological Innovation, Resource Allocation, and Growth," *Quarterly Journal of Economics*, 2017, *132* (2), 665–712.

Laffont, Jean-Jacques and Jean Tirole, *A Theory of Incentives in Procurement and Regulation*, Cambridge, MA: MIT Press, 1993.

Medicare Payment Advisory Commission, "Report to the Congress: Medicare Payment Policy," Technical Report, MedPAC, Washington, DC March 1999.

_ , "Report to the Congress: Medicare Payment Policy," Technical Report, MedPAC, Washington, DC March 2002.

Newhouse, Joseph P., "Medical Care Costs: How Much Welfare Loss?," *Journal of Economic Perspectives*, 1992, *6* (3), 3–21.

Rogers, Parker, "Regulating the Innovators: Approval Costs and Innovation in Medical Technologies," Working Paper, Indiana University 2026.

Rogerson, William P., "Economic Incentives and the Defense Procurement Process," *Journal of Economic Perspectives*, 1994, *8* (4), 65–90.

Sheingold, Steven H. and Elizabeth Richter, "Are PPS Payments Adequate? Issues for Updating and Assessing Rates," *Health Care Financing Review*, 1992, *14* (2), 165–175.

Steiner, Claudia A., Eric B. Bass, Mark A. Talamini, Henry A. Pitt, and Earl P. Steinberg, "Surgical Rates and Operative Mortality for Open and Laparoscopic Cholecystectomy in Maryland," *New England Journal of Medicine*, 1994, *330* (6), 403–408.

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

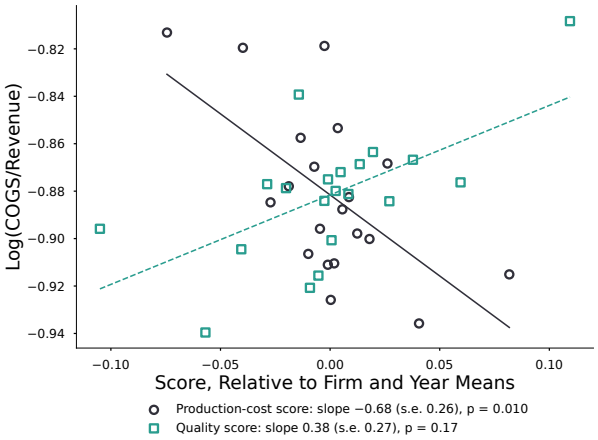
to Albert Ma, Ching, "Health Care Payment Systems: Cost and Quality Incentives," *Journal of Economics & Management Strategy*, 1994, *3* (1), 93–112.

Weisbrod, Burton A., "The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care, and Cost Containment," *Journal of Economic Literature*, 1991, *29* (2), 523–552.

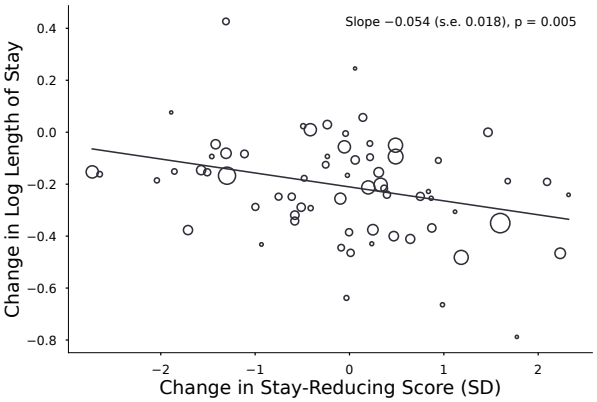
Figures and Tables

Figure 1: Patent Scores and Realized Costs

(a) Production-Cost Score and Firm Costs



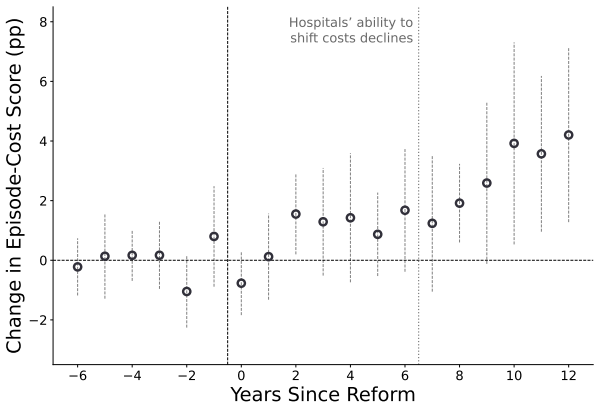
(b) Stay-Reducing Score and Hospital Stays



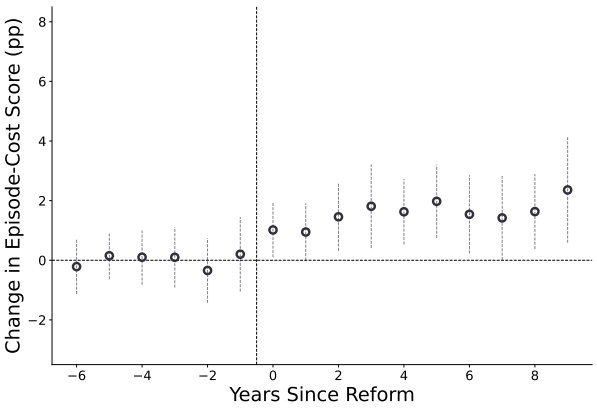
Note: Each panel plots an outcome against a patent score, both net of the regression's controls, with the fitted line, its slope, standard error, and p-value. *Panel (a)*: publicly traded firms holding medical-device patents (575 firms, 6,037 firm-years). Each marker is the mean of the log ratio of cost of goods sold to revenue within one of 20 equal-weight bins of the firm's mean production-cost score (circles) or quality score (squares) on a 0–1 scale, averaged over its patents filed from three years before to three years after the year (drug and cosmetic classes excluded; at least two patents). Firm and year fixed effects; observations weighted by the share of the firm's patents that are scored; standard errors clustered by firm. *Panel (b)*: 67 inpatient procedure categories of the AHRQ Clinical Classifications Software with at least 30 linked patents filed in each of 1976–82 and 1983–90 (National Hospital Discharge Survey). Each circle is a category, with area proportional to its 1979–82 discharges; the vertical axis is the change in mean log length of stay from 1979–82 to 1984–90, and the horizontal axis is the change in the stay-reducing score of the category's linked patents from 1976–82 to 1983–90, in standard deviations (2.7 percentage points). Both net of the category's 1979–82 log length of stay and its inpatient payment exposure; weighted by 1979–82 discharges; heteroskedasticity-robust standard errors. The same regression on the change from 1979–80 to 1981–82, before the reform, gives −0.024 (s.e. 0.010).

Figure 2: Episode-Cost Invention After Two Payment Reforms

(a) 1983 Inpatient Prospective Payment

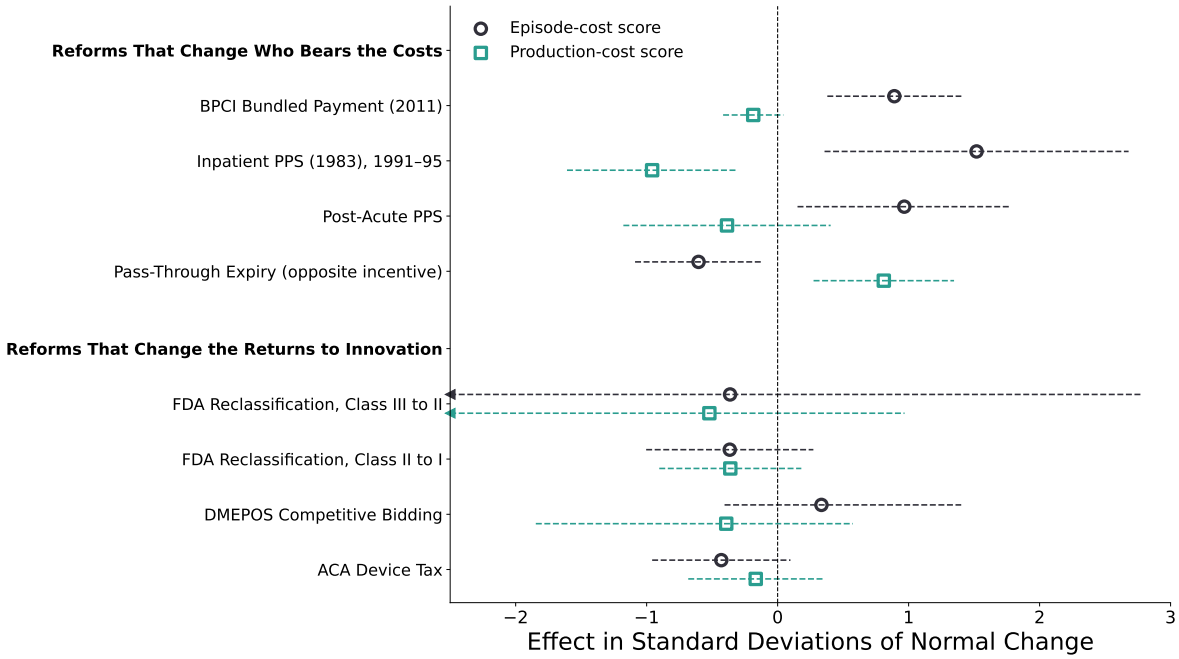


(b) BPCI Bundled Payment



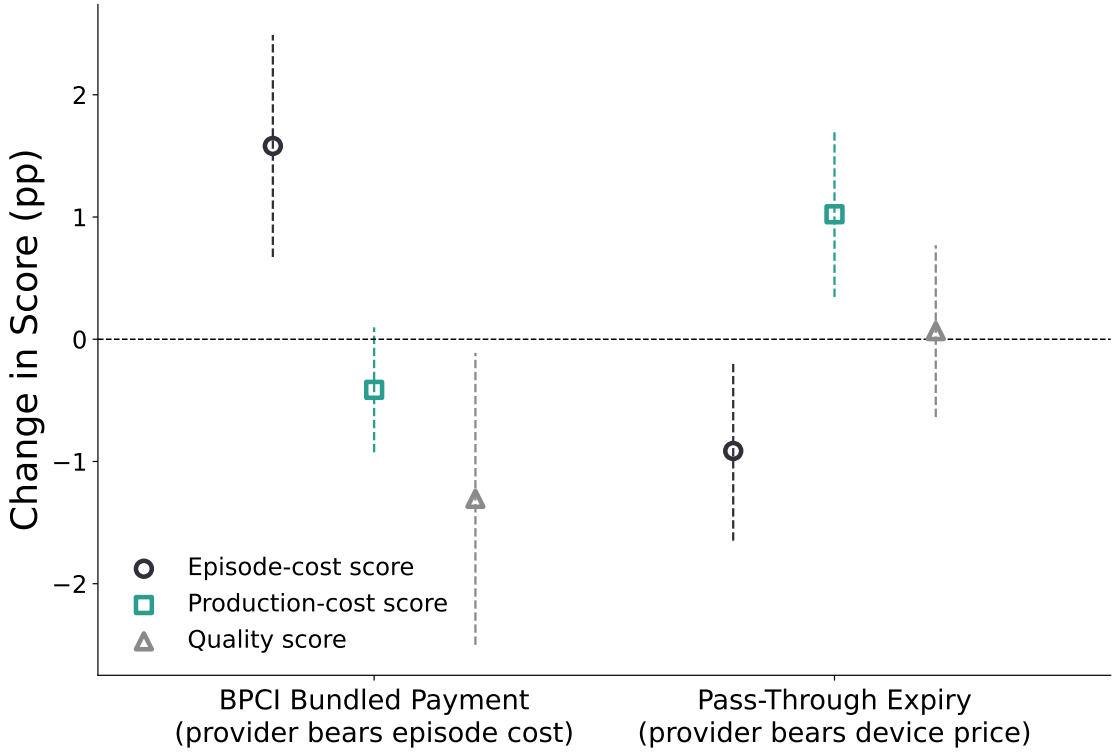
Note: Each point is the difference in the episode-cost score of device patents between exposed and comparison fields in a filing year, relative to the average difference over the six years before the reform; lines are 95% confidence intervals. Year 0 is 1983 in *panel (a)* and 2011 in *panel (b)*. The dotted line in *panel (a)* marks 1990, when hospitals' ability to shift Medicare's payment cuts onto private insurers declined (Cutler, 1998) and Medicare inpatient margins fell below zero. Both panels use the same vertical scale. *Panel (a)* compares 27 inpatient hospital-technology fields with 30 device fields used mainly outside inpatient care (drug fields and dental comparison fields excluded), filing years 1977–95, with field and year fixed effects and standard errors clustered by field (N = 32,328 patents). The response is 3.31 percentage points over 1991–95 (s.e. 1.29) and 1.03 over 1983–90 (s.e. 0.72); pre-reform coefficients are jointly insignificant ($p = 0.15$). *Panel (b)* compares patents with any procedure link to a code covered by a BPCI episode with other patents linked to device-intensive surgical procedures, with technology-class and year fixed effects and standard errors clustered by class (N = 74,780). The response is 1.58 percentage points (s.e. 0.46); pre-reform coefficients are jointly insignificant ($p = 0.97$); with all comparison patents it is 1.29 (s.e. 0.44). The 1991–95 response is 1.5 times, and the BPCI response 0.9 times, the standard deviation of the change in the score across fields in years without a reform (Figure 3). Section 4 defines the score; Appendix Table A.7 reports alternative comparisons.

Figure 3: Effects on the Direction of Invention by Reform Type



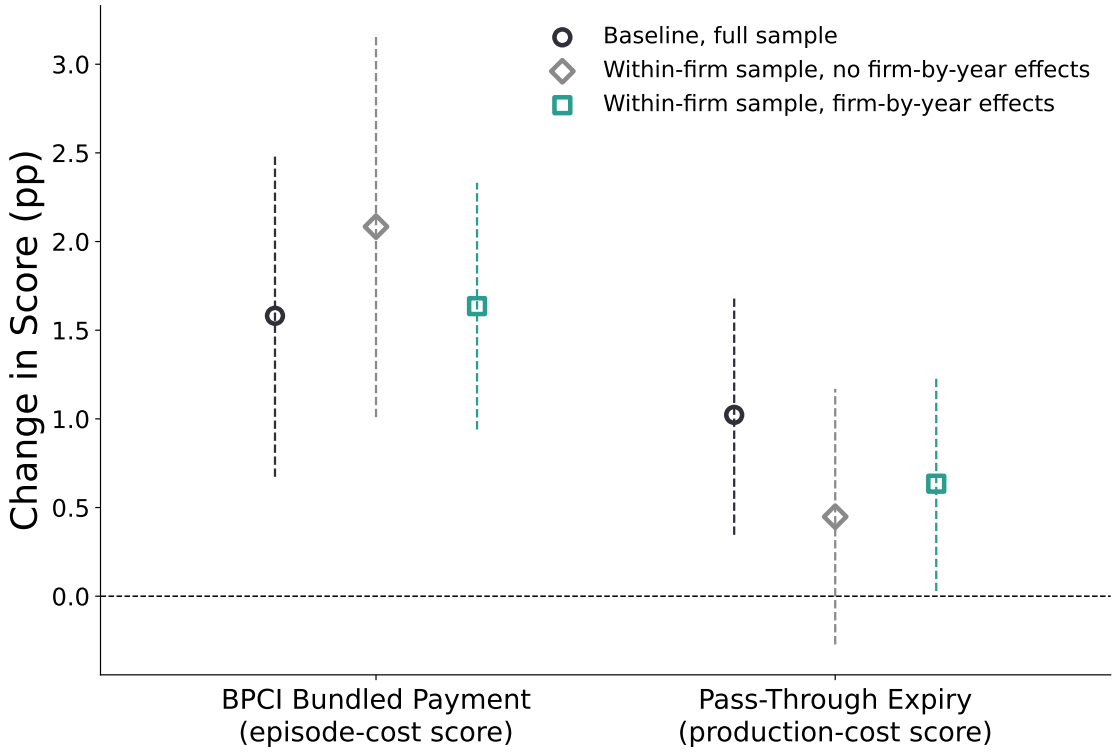
Note: Each marker is a reform's effect on a patent-text score divided by the standard deviation of normal changes in that score among comparison technologies over windows of the same length (Section 7.4); bars are 95% confidence intervals for the effect, divided by the same standard deviation. Circles are the episode-cost score and squares the production-cost score. Each reform that changes who bears the costs raises the score for the cost the provider now bears, and its point estimate for the other score is negative. Estimates are in Tables 4 and 5; the post-acute row pools the first payment reform in each post-acute setting. Intervals are cluster-robust, except FDA Class III to II and DMEPOS, which invert a wild cluster bootstrap test with the null imposed (12 treated product codes; 13 of 47 categories). The arrow marks an interval that extends beyond the axis. Appendix Figures A.2 and A.3 plot the post-acute and pass-through event studies.

Figure 4: Each Payment Rule Moves Invention Toward the Cost the Provider Bears



Note: Each marker is a difference-in-differences estimate of the change in a patent score, in percentage points; lines are 95% confidence intervals. *BPCI*: the provider bears the cost of the episode of care. The estimate compares patents with any procedure link to a code covered by a BPCI episode with other patents linked to device-intensive surgical procedures, after 2011, with technology-class and year fixed effects; standard errors clustered by class (N = 74,780). Episode cost 1.58 (s.e. 0.46), production cost −0.41 (0.26), quality −1.31 (0.61). *Pass-through expiry*: the provider bears the device's price once Medicare's separate payment for it ends. The estimate compares patents linked to the 80 device categories that received pass-through payment, after the payment ends, with the same categories while it was paid, relative to patents linked to other codes; stacked by start year, with code-by-stack and year-by-stack fixed effects and standard errors clustered by code. Episode cost −0.91 (0.37), production cost 1.02 (0.35), quality 0.07 (0.36). Section 4 defines the scores.

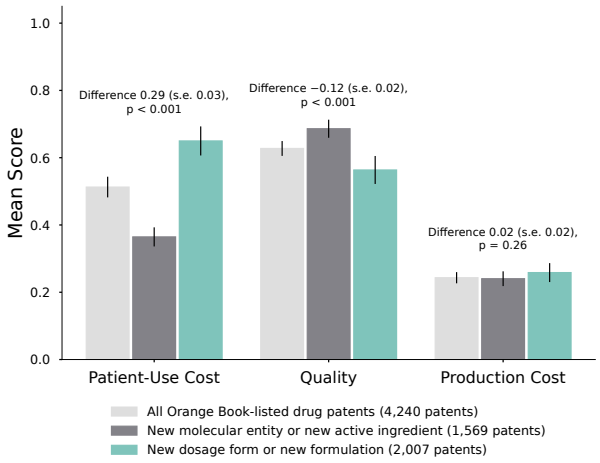
Figure 5: The Response Holds Within Firm and Year



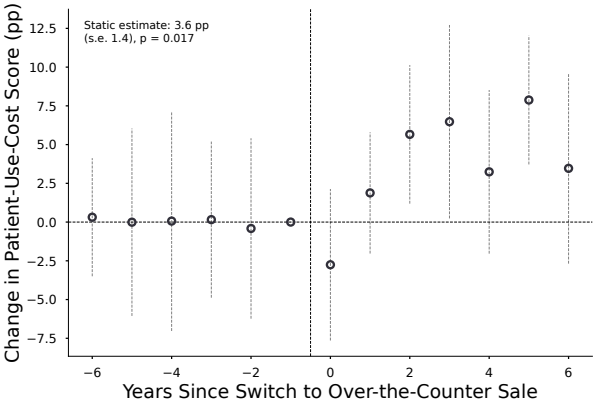
Note: For each reform, the baseline estimate on the full sample (circles), the estimate on the within-firm sample without firm-by-year fixed effects (diamonds), and the estimate on the same sample comparing patents of the same firm filed in the same year (squares), in percentage points; lines are 95% confidence intervals. The within-firm sample keeps patents of firms with both exposed and comparison patents in a filing year. *BPCI*: episode-cost score, incumbent firms, technology-class fixed effects; standard errors clustered by firm and technology class. Baseline 1.58 (s.e. 0.46, $N = 74{,}780$); within-firm sample 2.08 (0.55) and 1.64 (0.35), both $N = 44{,}285$. *Pass-through expiry*: production-cost score, exposed minus shielded years, standard errors clustered by code. Baseline 1.02 (0.35); within-firm sample 0.45 (0.37, $N = 78{,}782$) and 0.63 (0.31, $N = 77{,}088$). Only BPCI and pass-through expiry are split, because a comparison within firm and year needs firms that file both exposed and comparison patents in the same year. Such firm-years hold 67 percent of exposed BPCI patents, against 30 percent in the post-acute design and 7 percent in the 1983 design, where 40 percent of patents also have no company assignee. On the post-acute within-firm sample the estimate is 0.30 percentage points (s.e. 0.64) before firm effects, against 1.11 on the full sample, so that sample cannot separate change within firms from change across them. The reforms that change only the returns to innovation show no response to split.

Figure 6: The Drug Measure and Prescription-to-OTC Switches

(a) Drug Patents by FDA Approval Type



(b) Prescription-to-OTC Switches



Note: *Panel (a)*: mean score, by channel, of the drug patents that manufacturers list in FDA's Orange Book, by FDA's classification of the application they protect: new molecular entities or new active ingredients (1,569 patents; Drugs@FDA chemical types 1 and 2) and new dosage forms or new formulations (2,007 patents; types 3 and 5). Grey bars are all 4,240 listed patents; patents listed on both kinds of application are excluded from the two groups. Lines are 95% confidence intervals clustered by application. Each difference is the mean score of new-dosage-form patents minus that of new-molecule patents filed in the same year, from a regression of the score on an indicator for new-dosage-form patents with filing-year fixed effects; standard errors clustered by application. A new dosage form changes how the patient takes the drug, which the patient-use-cost score measures, and new molecules are the main source of new therapeutic benefit, which the quality score measures. Neither approval type has an obvious claim on lower manufacturing cost, and the production-cost score does not differ between them. *Panel (b)*: event-study coefficients for the patient-use-cost score of patents linked to the 24 molecules FDA switched from prescription to OTC sale, relative to five matched never-switched molecules per switch, by years since FDA approved OTC sale (reference: the year before). Stacked by switch, with molecule-by-stack and year-by-stack fixed effects, weighted by patents; 95% confidence intervals clustered by stack. The static estimate is the difference-in-differences coefficient on all post-switch years from the same stacked regression, with p-values from a t distribution with 23 degrees of freedom.

Table 1: Predicted and Estimated Direction of Invention, by Reform

Reform	Direction		What the reform changes
	Predicted	Estimated	
A. Reforms That Change Who Bears a Cost			
Inpatient PPS (1983)†	up	up	provider bears episode cost
BPCI (2011)	up	up	provider bears episode cost
Post-acute PPS	up	up	provider bears episode cost
Pass-through expiry‡	up	up	provider bears device price
B. Reforms That Change the Size of the Payoff			
FDA reclassification III→II	0	≈0	lower entry cost
FDA reclassification II→I	0	0	lower entry cost
DMEPOS competitive bidding	0	0	lower price
ACA device tax	0	0	tax on revenue

Note: Direction is the outcome each design uses in Tables 4 (Panel A) and 5 (Panel B): the episode-cost score in Panel A, except the production-cost score for the row marked ‡, and the production-cost score in Panel B. up marks a shift toward cost reduction, 0 no change, and ≈0 an estimate too imprecise to distinguish a shift from no change. Predicted directions follow from Section 3; for Panel B, they hold under the benchmark in Appendix B, and the ACA prediction assumes manufacturers bear the tax. Estimates are in Tables 4 (Panel A) and 5 (Panel B). †The estimate is the response in 1991–95, relative to 1977–82. ‡Device categories whose separate payment ended, against those still paid separately; the production-cost score rises by 1.02 percentage points and the episode-cost score falls by 0.91.

Table 2: Examples of Patents in Each Direction

Patent title (filing year)	What the invention does	Score		
		Production cost	Episode cost	Quality
Quality-expanding				
Flexible circuit electrode array embedded in a cured body (2013)	A flexible electrode array for stimulating nerves, used to create artificial vision or hearing.	0.08	0.20	0.89
Ultrasound imaging system memory architecture (2020)	An ultrasound system that stores raw echo data so images can be reprocessed to show detail not visible in the original image.	0.17	0.16	0.86
Episode-cost-reducing				
Cardiac valve procedure methods and devices (2004)	Devices for replacing a heart valve through the blood vessels, without putting the patient on cardiac bypass.	0.06	0.90	0.27
Devices and methods for percutaneous access, hemostasis, and closure (2014)	Devices that control bleeding and close the opening after an organ is reached through the skin.	0.07	0.87	0.23
Production-cost-reducing				
Method for making containers having ports (1985)	A method that molds a fluid container's port onto its plastic pouch in one step, removing a separate heat-sealing operation to lower manufacturing cost.	0.92	0.20	0.15
Two-ply sterilization wrap and method for sterilizing an article (1998)	A sterilization wrap folded from a single sheet, giving two layers of protection while staying inexpensive to manufacture.	0.88	0.24	0.13

Note: Patents come from the validation sample, which the classifiers never saw in training, so the scores show how the measure classifies new patents rather than ones it was fit to; consumer products are excluded. For each direction, I keep patents whose other two scores are both below 0.35, rank them by the direction's score, and show two of the top five whose text states the direction most plainly. Scores are the classifier's probabilities; the episode-cost score averages the complication, length-of-stay, and site-of-care scores. The descriptions paraphrase each patent's abstract and summary of the invention, the text the classifiers read.

Table 3: Summary Statistics

Variable	N	Mean	SD	Range
I. Device Patents				
Production-Cost Score	411,651	0.215	0.137	[0.019, 0.977]
Episode-Cost Score	411,651	0.316	0.122	[0.052, 0.921]
Quality-Expanding Score	411,651	0.504	0.196	[0.026, 0.955]
II. Bundled Payment (BPCI)				
Episode-Cost Score	20,921	0.377	0.137	[0.096, 0.921]
Production-Cost Score	20,921	0.152	0.093	[0.020, 0.871]
Exposed (Any Procedure Link to a BPCI Episode)	20,921	0.145	0.353	[0.000, 1.000]
Share of Procedure Links in BPCI Episodes, Exposed Patents	3,043	0.408	0.224	[0.186, 1.000]
III. Inpatient PPS (1983)				
Episode-Cost Score†	4,924	0.334	0.130	[0.066, 0.833]
Production-Cost Score	4,924	0.312	0.146	[0.057, 0.947]
Inpatient Technology Field	4,924	0.683	0.465	[0.000, 1.000]
IV. Post-Acute PPS				
Episode-Cost Score‡	53,351	0.301	0.129	[0.054, 0.932]
Production-Cost Score	53,351	0.215	0.142	[0.024, 0.953]
Treated Group	53,351	0.247	0.431	[0.000, 1.000]
V. Pass-Through Expiry				
Production-Cost Score	137,054	0.191	0.097	[0.021, 0.933]
Episode-Cost Score	137,054	0.318	0.106	[0.069, 0.909]
VI. Prescription-to-OTC Drugs				
Patient-Use-Cost Score	159,361	0.325	0.207	[0.037, 0.991]
Production-Cost Score	159,361	0.231	0.186	[0.026, 0.977]
Quality-Expanding Score	159,361	0.744	0.183	[0.039, 0.969]

Note: Scores are classifier probabilities (Section 4); the episode-cost score averages the complication, length-of-stay, and site-of-care scores. Panels II–V report pre-reform means, standard deviations, and ranges for the estimation samples behind the baseline estimates in Table 4, exposed and comparison patents together (that table reports exposed patents alone): BPCI filings 2005–2010, 1983 filings 1977–1982, the years before each post-acute reform, and the years before pass-through starts. Panels I (device patents issued 1976–2024) and VI are not tied to a reform. Panel II: 641 technology classes; a patent is exposed if any of its procedure links falls on a code covered by a BPCI episode, and the last row reports, for exposed patents, the share of their procedure links that do. Panel III: the 26 of 27 inpatient technology fields and 26 of 30 comparison device fields with patents filed before 1983. Panel IV: one row per patent and reform stack (31,091 distinct patents) in the 24 of 26 treated groups with pre-reform patents and 49 matched comparison groups. Panel V: code-by-year cells in each of the 11 reform stacks (comparison codes enter every stack), weighted by patents. Panel VI: A61K drug patents (Section 8). †Complications and length of stay only (inpatient reform). ‡The channels each setting's payment rule rewards: all three for home health; complications and length of stay for rehabilitation and long-term care.

Table 4: Reforms That Change Who Bears the Costs

	Pre-Reform Mean (%)	Baseline	Baseline in SDs of Normal Change	Alternative Comparison	Within Firm
	(1)	(2)	(3)	(4)	(5)
A. Bundled Payment (BPCI, 2011)					
Episode-Cost Score	36.4	1.58***	0.89	1.29**	1.64***
		(0.46)		(0.44)	(0.35)
Pre-Trend p		0.97		0.75	0.44
Sample Size		74,780		128,959	44,285
B. Inpatient Prospective Payment (1983)					
Episode-Cost Score, 1983–90	37.6	1.03	–	0.95	–
		(0.72)		(0.65)	
Episode-Cost Score, 1991–95		3.31*	1.52	2.47*	–
		(1.29)		(1.17)	
Pre-Trend p		0.15		0.50	
Sample Size		32,328		58,484	
C. Post-Acute Prospective Payment					
Episode-Cost Score	27.6	1.11*	0.97	0.98+	–
		(0.48)		(0.50)	
Pre-Trend p		1.00		–	
Sample Size		131,525		–	
D. Pass-Through Expiry					
Production-Cost Score	16.7	1.02**	0.81	–	0.63*
		(0.35)			(0.31)
Pre-Trend p		0.93			0.90
Episode-Cost Score	37.4	−0.91*	−0.60	–	−0.54
		(0.37)			(0.36)
Pre-Trend p		0.78			0.87
Sample Size		446,154			77,088

Note: Each estimate is the reform's effect on a patent-text score, in percentage points; standard errors, clustered on the design's treatment unit (Table A.5), are in parentheses. The episode-cost score is the probability that an invention lowers the cost of an episode of care (fewer complications, shorter stays, or a cheaper setting), averaged over the channels the reform's payment rule rewards; the production-cost score is the probability that it lowers the device's manufacturing cost. Column (1) is the mean score of exposed patents before the reform. Column (3) divides column (2) by the standard deviation of normal changes in the score among comparison technologies (Section 7.4). Column (4) compares BPCI patents with every procedure-linked patent, the 1983 fields with all other fields filed 1976–95, and post-acute groups with a broader matched pool. Column (5) compares patents of the same firm filed in the same year. A BPCI patent is exposed if any of its procedure links falls on a code covered by a BPCI episode. Pre-trend p tests that the event-study coefficients before the reform (for pass-through, before pass-through starts) are jointly zero. Sample size counts patents, except the pass-through baseline, which counts stacked code-by-year cells. The production-cost score does not change when pass-through begins (−0.05, s.e. 0.30). +, *, **, and *** denote statistical significance at the 0.10, 0.05, 0.01, and 0.001 levels, respectively.

Table 5: Reforms That Change the Returns to Innovation

	Pre-Reform Mean (%) (1)	Baseline (2)	Baseline in SDs of Normal Change (3)	Pre-Trend p (4)	Sample Size (5)
FDA reclassification, Class III to II					
Production-Cost Score	19.1	−0.96	−0.52	0.91	1,118
		(1.87)			
Episode-Cost Score	21.6	−0.57	−0.36	0.07	1,118
		(2.32)			
FDA reclassification, Class II to I					
Production-Cost Score	40.1	−0.76	−0.36	0.51	10,835
		(0.58)			
Episode-Cost Score	20.3	−0.51	−0.37	0.05	10,835
		(0.45)			
DMEPOS competitive bidding					
Production-Cost Score	21.3	−0.83	−0.39	0.95	8,613
		(0.70)			
Episode-Cost Score	34.1	0.47	0.33	0.38	8,613
		(0.36)			
ACA device tax					
Production-Cost Score	25.2	−0.22	−0.17	0.85	6,301
		(0.35)			
Episode-Cost Score	32.6	−0.39	−0.43	0.39	6,301
		(0.24)			

Note: Each estimate is the reform's effect on a patent-text score, in percentage points, with standard errors in parentheses. The production-cost score is the probability that an invention lowers the device's manufacturing cost; the episode-cost score is the probability that it lowers the cost of an episode of care (fewer complications, shorter stays, or a cheaper setting). Column (1) is the mean score of treated product codes or categories before the reform. Column (3) divides column (2) by the standard deviation of normal changes in the score among comparison technologies (Section 7.4). Pre-trend p tests that the event-study coefficients before the reform (up to six years; the year before is the reference) are jointly zero. Sample size counts product-code-by-year (FDA, device tax) or category-by-year (DMEPOS) cells. Standard errors are clustered by product code or category. The Class III-to-II and DMEPOS rows have few treated units (12 product codes; 13 of 47 categories), so their significance levels and pre-trend tests use a wild cluster bootstrap. For the FDA reforms, the episode-cost pre-trend tests reject at the 10 percent level because the year before the reform lies off an otherwise flat series; the event studies show no break at the reform (Appendix Figure A.1). Designs are in Table A.5. +, *, **, and *** denote statistical significance at the 0.10, 0.05, 0.01, and 0.001 levels, respectively.

A Institutional Detail on the Payment Reforms

This appendix describes the payment rules, timing, and affected technologies for the reforms used in the analysis. I also provide details on the construction of treatment and comparison groups and the alternative specifications. Section 6.2 discusses the 1983 inpatient prospective payment system. Appendix Table A.1 lists related reforms I considered but excluded.

Entry deregulation (FDA reclassification). FDA reclassification moves device types to less restrictive regulatory pathways, reducing the costs of entering the market. I use the treated device types, comparison groups, patent links, and stacked difference-in-differences estimators from Rogers (2026). Reclassification substantially increases patenting but does not produce a significant change in the direction of invention.

For the twelve FDA-initiated Class III-to-II reclassifications, the production-cost score changes by −0.96 percentage points (s.e. 1.87), while the episode-cost score changes by −0.57 (s.e. 2.32). For the 129 Class II-to-I reclassifications in 1996 and 1998 with linked patents, the corresponding estimates are −0.76 (s.e. 0.58) and −0.51 (s.e. 0.45). The latter confidence intervals rule out increases larger than approximately 0.4 percentage points in either measure. Pre-reform coefficients for production cost are jointly insignificant in both designs (Table 5).[54] These results are consistent with reclassification changing the overall returns to invention without changing the relative returns to reducing production or episode costs.

The 1990–1991 margin squeeze and 1992 capital prospective payment. Medicare introduced prospective payment for hospital capital costs in 1992, shortly after hospitals' operating margins turned negative. The capital reform gradually incorporated equipment and other capital costs into the DRG payment over ten years. It therefore provides a possible alternative explanation for the larger shift in invention after 1990–1991.

The evidence does not suggest that capital prospective payment explains the result. First, the increase occurs in inventions that reduce complications and length of stay, rather than those that reduce device production costs. The latter response is more directly associated with reforms that make hospitals responsible for device purchase prices, as in the outpatient pass-through analysis. Second, the increase extends across inpatient technology fields rather than concentrating in durable equipment. Third, the response begins after 1983, well before capital prospective payment, and the larger response after 1990 remains when I divide the two periods at any year from 1987 to 1994 (Section 6.2). The capital reform was also phased in gradually, whereas the estimated increase in episode-cost invention occurs over a much shorter period.

I further test for a change in invention in durable-equipment fields around the introduction of capital prospective payment. I find no significant break in 1992, although the gradual phase-in limits the power of this test. Taken together, the timing, affected technologies, and direction of invention are more consistent with the incentives created by inpatient prospective

[54] I use the matched samples, patent links, and stacked estimators from Rogers (2026), including event-by-product-code and event-by-year fixed effects. Product-code-year observations are weighted by their number of linked patents. The Class III-to-II design includes a two-year anticipation window and uses a wild cluster bootstrap ($p = 0.54$ for production cost). The Class II-to-I design uses a one-year anticipation window. Applying the original code to patent counts reproduces the published estimates.

payment than with the later incorporation of capital costs.

Bundled payment (BPCI). Bundled payments make providers financially responsible for spending across an episode of care, including hospitalization, physician services, and post-acute care. Under the principal BPCI models, Medicare continues to reimburse individual services and subsequently reconciles total episode spending against a target price. Participating providers retain part of any savings below the target and bear part of the spending above it. The reform therefore changes their incentives to reduce costs across services that were previously reimbursed separately.

The Affordable Care Act authorized Medicare's bundled-payment initiative in March 2010 but did not specify the procedures it would cover. CMS identified 48 clinical episodes, including lower-extremity joint replacement, in its August 2011 solicitation. Providers began bearing financial risk in October 2013, with participation phased in across cohorts. I date the reform to August 2011, when the covered episodes became known and manufacturers could respond. A patent is classified as exposed if at least one of its five closest procedure-code links is covered by a BPCI episode (Section 4). Joint-replacement episodes involved approximately \$27,000–31,000 in spending (Dummit et al., 2016).

Medicare subsequently introduced the Comprehensive Care for Joint Replacement (CJR) program, announced in 2015 and implemented in 2016. Unlike BPCI, CJR required hospitals in selected metropolitan areas to accept episode-based payments for lower-extremity joint replacements. Because these procedures were already covered by BPCI, I do not estimate CJR separately.

Post-acute prospective payment. The Balanced Budget Act of 1997 and the Balanced Budget Refinement Act of 1999 replaced cost-based reimbursement with prospective payment in four post-acute settings. Skilled nursing facilities received fixed payments per day, including routine services, ancillary services, and capital costs. Home health agencies received payments per episode, including medical supplies. Inpatient rehabilitation facilities and long-term care hospitals received payments per discharge. These rules changed which costs each provider was responsible for, but differed in the incentives they created to shorten stays or reduce the use of care.

I include home health, inpatient rehabilitation, and long-term care hospitals in the analysis. Their payment rules reward reductions in the costs of an episode or discharge. I exclude skilled nursing facilities because per-day reimbursement does not directly reward shorter stays, although facilities may still benefit from reducing costs within a day. I also exclude inpatient psychiatric facilities, which moved to prospective payment in 2005 but are not a post-acute setting.

I date each reform to the first proposed or interim final rule specifying its payment design. The dates are 2000 for home health (October 1999 proposed rule), 2001 for inpatient rehabilitation (November 2000 proposed rule), and 2002 for long-term care hospitals (March 2002 proposed rule). Rules published late in the calendar year are assigned to the following year. Using implementation instead gives an episode-cost estimate of 1.14 percentage points (s.e. 0.49), compared with 1.11 under the announcement dates, with jointly insignificant pre-reform coefficients ($p = 0.80$).[55]

[55]Implementation began in October 2000 for home health, January 2002 for inpatient rehabilitation, and

I identify affected technologies using the supplies and equipment covered by each setting's new payment rule. For each setting, I first identify the relevant product categories and then assign them to CPC main groups. A technology is treated only if the payment system covers its costs. For example, home health episode payments included dressings and catheters but excluded durable equipment such as wheelchairs, nebulizers, infusion pumps, and home monitors, which Medicare continued to reimburse separately.

I assign each technology group to the first reform that covers it. Home health includes dressings and catheters beginning in 2000. Inpatient rehabilitation includes therapy equipment, orthoses, wheelchairs, and walking aids beginning in 2001. Long-term care hospitals include beds, feeding tubes, monitoring equipment, respiratory devices, and disinfection technologies beginning in 2002. The outcome for each setting follows the costs rewarded by its payment rule (Section 5).

Some technologies used in these settings are not included among the treated groups, including bed-pans, nebulizers, wound-drainage equipment, medical tubing, and medication dispensers. Because these technologies may also have been affected, I exclude them from the comparison group, along with dental technologies and funeral equipment. Other technologies used primarily outside post-acute care, such as ambulances, remain eligible as comparisons.

The estimates are not sensitive to most of these choices. Removing the coverage restriction, excluding dressing-material technologies, or excluding home health produces estimates between 1.03 and 1.36 percentage points (Appendix Table A.11). Using the broader pool of acute-care technologies, including the excluded groups, gives 0.98 percentage points (s.e. 0.50). By contrast, treating skilled nursing facilities as exposed reduces the estimate to 0.50 percentage points (s.e. 0.52), and the pre-reform coefficients become jointly significant ($p = 0.05$). This result is consistent with per-day reimbursement providing different incentives from payments fixed per episode or discharge.

Outpatient prospective payment and device pass-through. Medicare introduced the Outpatient Prospective Payment System (OPPS) in August 2000, replacing cost-based reimbursement with fixed payments for outpatient procedures. These payments, organized into ambulatory payment classifications (APCs), generally include the cost of devices used during the procedure. Hospitals therefore bear the cost of purchasing those devices.

To avoid discouraging the adoption of costly new technologies, Congress created transitional pass-through payments for qualifying devices under the Balanced Budget Refinement Act of 1999 (Section 201). These payments provide additional reimbursement for two to three years, covering the amount by which the device's cost exceeds the payment already included in the APC rate. While the pass-through is in effect, hospitals are partly insulated from the device's purchase price. When it expires, the device is included in the fixed procedure payment, increasing the hospital's incentive to purchase less expensive devices. CMS explicitly identifies this incentive as a rationale for packaging device costs into procedure payments.[56]

October 2002 for long-term care hospitals. Skilled nursing prospective payment began in 1998 under an interim final rule published in May of that year.

[56]See the CMS OPPS final rules and the *Medicare Claims Processing Manual*, Chapter 4, on transitional pass-through payments and packaging. The Balanced Budget Act of 1997 authorized outpatient prospective payment, and the 1999 legislation added device pass-through payments before the system began in August 2000.

I obtain the start and expiration dates for each device category from CMS. The expiration date is determined when pass-through status is granted, making the change in reimbursement foreseeable. Nevertheless, production-cost invention does not increase when separate reimbursement begins and rises after it expires. The result also holds when I compare categories at the same time since pass-through began, using variation between devices receiving two and three years of separate reimbursement (Section 6.3). Of the 86 device categories granted pass-through status between 2000 and 2018 with linked patents, 80 enter the estimation sample.

Medicare competitive bidding for durable medical equipment. Medicare historically reimbursed durable medical equipment, prosthetics, orthotics, and supplies (DMEPOS) through administratively determined fee schedules. Following the Medicare Modernization Act of 2003, Medicare gradually replaced these rates with prices determined through competitive bidding in selected markets, beginning in 2011. The resulting price reductions lowered the returns to selling devices in affected categories.

Unlike prospective payment, competitive bidding did not change which party bore the marginal costs of producing or supplying the device. Suppliers were responsible for their own costs under both the fee schedule and competitive bidding. The reform therefore changed reimbursement levels without directly changing the relative incentives to reduce production costs rather than improve quality.

I use the 13 treated categories among 47 DME categories, the 1996–2016 sample, patent links, and stacked estimator from Ji and Rogers (2024). Treatment begins in the year each category's inclusion was announced (2004, 2007, 2011, or 2012). The corresponding estimates using patent counts reproduce those in the earlier paper.

The medical-device excise tax. The Affordable Care Act imposed a 2.3 percent excise tax on manufacturers' sales of taxable medical devices. The tax took effect in 2013 and was suspended after 2015. It applied to most medical devices but exempted products commonly purchased directly by consumers at retail, including eyeglasses and hearing aids. The tax reduced the net returns to device sales without changing which costs providers bore.

I compare 515 taxable product codes with 159 exempt codes, matched on their pre-reform production-cost scores and patenting volume during 2005–2009. I date the reform to its enactment in 2010, when manufacturers could anticipate the tax. The estimated effect remains statistically insignificant when I instead use its implementation in 2013.

Matching is important for this comparison. Without matching, I estimate an increase of 0.61 percentage points (s.e. 0.29), using 2,892 taxable and 304 exempt codes. However, taxable codes begin with substantially lower production-cost scores than exempt codes (0.17 versus 0.27), and among taxable codes those with lower initial scores subsequently increase more. Matching on pre-reform levels addresses this convergence.

Prescription-to-OTC switches. When the FDA allows a prescription drug to be sold over the counter, patients can purchase it without obtaining a prescription and generally pay for it themselves. The switch therefore changes both who chooses the drug and who bears its cost.

I use a curated list of FDA prescription-to-OTC switches, dated to the year of approval. The sample contains 24 molecules with sufficient patenting within six years before and after

approval. I link patents to molecules using Orange Book drug applications and retain patent-to-drug links with cosine similarity of at least 0.45. Because the main classifier is trained on medical-device patents, I construct a separate measure of drug inventions that reduce patients' costs of use, following the same classification approach (Appendix C).

I estimate a stacked difference-in-differences design, matching each switched molecule to five never-switched molecules using Mahalanobis distance. Standard errors are clustered by matched event stack. Treatment is defined at the ingredient-form level, and patents linked to multiple molecules within a stack are assigned to the most similar molecule. The patient-use-cost score increases by 3.6 percentage points after approval (s.e. 1.4, $p = 0.017$), with no significant pre-reform trend. Dating switches three or five years earlier gives smaller, statistically insignificant estimates, while using earlier announcement dates and rematching controls produces an estimate close to zero. These timing checks indicate that the response is concentrated after FDA approval rather than the earlier signals that a switch might occur.

Who chooses and pays for medical devices. I classify the 1,052 FDA product codes in the patent sample according to who typically chooses the device and who pays for it. The classifications describe typical U.S. use during 2000–2019 for commercially insured patients younger than 65.

I first use a language model to classify each device using its FDA product name, review panel, regulatory class, regulation text, and definition. A second model independently classifies every code initially identified as patient-chosen or jointly chosen, along with 60 randomly selected codes. The models agree on 87 percent of the random sample for both the choice and payment classifications. I then manually review all patient-chosen codes with at least partial insurance coverage and 20 additional randomly selected codes.

Patients choose the specific product for 51 device codes, of which 47 are generally purchased out of pocket and four have mixed payment. None is usually covered by insurance. Clinicians choose 225 codes, hospitals and other facilities or suppliers choose 755, and decisions are shared for 21. Thus, among insured medical devices, the specific product is generally chosen by a physician, hospital, or supplier rather than the patient. This limited variation motivates the use of prescription-to-OTC switches to examine the mechanism when patients themselves become the decision-makers.

B Detailed Conceptual Framework and Proofs

This appendix presents the model of directed innovation summarized in Section 3. The model illustrates how payment systems affect the types of inventions manufacturers develop. It also distinguishes changes in the direction of invention from changes in the overall returns to R&D and derives predictions for the payment reforms studied in the paper. I begin with a manufacturer choosing between reducing the costs of care and improving quality, and then extend the model to distinguish reductions in episode costs from reductions in device production costs.

Model Preliminaries

Consider a representative medical device manufacturer that chooses how much R&D to direct toward two improvements: reducing the costs of care associated with its device and improving the device's quality. The manufacturer develops these improvements in response to the value they create for the decision-maker who chooses the device, usually a hospital or other health care provider. The model is characterized by the following parameters.

R&D Effort and Costs.—The manufacturer chooses effort $a_c \geq 0$ to reduce the cost of care and effort $a_q \geq 0$ to improve quality. For simplicity, one unit of effort reduces the relevant cost c or increases quality q by one unit. R&D costs are quadratic:

$$C(a_c, a_q) = \frac{k}{2}(a_c^2 + a_q^2),$$

where $k > 0$. The marginal cost of effort on each dimension is ka_j, so directing more R&D toward either improvement becomes increasingly costly. I assume the two types of effort have the same cost function to isolate differences in their returns. I define the direction of invention as the share of R&D effort directed toward reducing costs:

$$s \equiv \frac{a_c}{a_c + a_q}.$$

This is a theoretical allocation of R&D effort, rather than the share of patents classified as cost-reducing. The empirical measures capture the objectives described in patents and need not sum to one.

Payment Incentives.—The payment system determines how much of the relevant cost the decision-maker bears. Let $\theta \in [0, 1]$ denote this share. When a provider is reimbursed for an additional dollar of spending, its effective cost exposure is low (θ near zero). Under a fixed payment, the provider generally bears that additional dollar and retains the savings from reducing it (θ near one). Thus, a reform that increases θ makes technologies that lower the costs of care more valuable to the provider.

For example, under cost-based reimbursement, a hospital has relatively little financial incentive to develop or adopt technologies solely because they shorten a hospital stay. Under prospective payment, shortening the stay reduces the hospital's costs without generally reducing its reimbursement. The hospital therefore becomes more willing to pay for technologies that achieve this improvement.

Willingness to Pay.—The decision-maker values device quality and the costs of care associated with using the device. I write its willingness to pay as

$$\text{WTP} = v_q q - \theta c + \text{constant},$$

where $v_q > 0$ measures the value of quality. I assume that payment design changes the share of costs the decision-maker bears but does not directly change its valuation of quality, v_q. This assumption isolates the cost-exposure mechanism. The device's production cost does not enter the expression because the buyer does not directly observe or bear the manufacturer's production costs. I introduce this margin separately below.

Returns to Invention.—The manufacturer captures part of the increase in willingness to pay generated by its inventions. Let $\mu > 0$ denote the common scale of these returns, reflecting factors such as the expected number of units sold and the extent to which the manufacturer can capture improvements in the buyer's willingness to pay. For the baseline model, I assume that these factors affect the returns to both types of invention proportionally. This assumption allows me to distinguish changes in the overall returns to invention from changes in the relative value of cost reduction and quality improvement.

The Manufacturer's Problem

Reducing the cost of care by a_c increases the decision-maker's willingness to pay by θa_c. Improving quality by a_q increases it by $v_q a_q$. The manufacturer therefore chooses R&D effort to solve

$$\max_{a_c, a_q \geq 0} \quad \mu(\theta a_c + v_q a_q) - \frac{k}{2}(a_c^2 + a_q^2).$$

The first-order conditions imply

$$a_c^\star = \frac{\mu\theta}{k}, \qquad a_q^\star = \frac{\mu v_q}{k}.$$

Thus, effort on each margin depends on the value of that improvement to the buyer, the returns the manufacturer can capture, and the cost of R&D. The share of effort directed toward reducing costs is

$$s^\star = \frac{\theta}{\theta + v_q},$$

while total effort is

$$a_c^\star + a_q^\star = \frac{\mu(\theta + v_q)}{k}.$$

These expressions illustrate the main distinction in the model. Payment design determines how valuable cost reduction is relative to quality improvement. Other factors, such as the size of the market, can change the returns to both types of invention without changing their relative value.

Payment Design and the Direction of Invention

When the decision-maker bears more of the costs of care, reducing those costs becomes more valuable. The manufacturer responds by allocating a larger share of R&D toward cost reduction.

Proposition 1 (Residual claimancy shifts invention toward cost reduction)**.** *An increase in the share of costs borne by the decision-maker raises the share of R&D directed toward cost reduction:*

$$\frac{\partial s^\star}{\partial \theta} = \frac{v_q}{(\theta + v_q)^2} > 0.$$

The intuition is straightforward. A technology that saves the provider one dollar becomes more valuable when the provider retains more of that saving. This increases the manufacturer's return to developing such a technology relative to improving quality. In the model, greater cost exposure increases cost-reducing effort without reducing quality-improving effort, because the return to quality remains unchanged. The direction of invention changes because cost-reducing effort increases more.

A different prediction follows for changes that affect the overall returns to invention. Suppose the market becomes larger, increasing the expected sales of both types of improvements proportionally. The manufacturer then invests more in both cost reduction and quality, but the relative allocation of effort does not change. The same result holds when the marginal costs of both types of R&D change proportionally.

To show that this result does not depend on quadratic costs, suppose the cost of effort on each margin is

$$C_j(a_j) = \frac{k}{1 + 1/\varepsilon} a_j^{1+1/\varepsilon}, \qquad \varepsilon > 0,$$

where ε is the elasticity of effort with respect to its return. The quadratic case corresponds to $\varepsilon = 1$.

Proposition 2 (Common changes in the returns to invention affect its level, not its direction)**.** *If the returns to both types of invention change proportionally, and the cost functions have the same constant elasticity, the share of effort directed toward cost reduction is*

$$s^\star = \frac{\theta^\varepsilon}{\theta^\varepsilon + v_q^\varepsilon}.$$

It is independent of the common return parameter μ, while total effort is

$$a_c^\star + a_q^\star = \left(\frac{\mu}{k}\right)^\varepsilon \left(\theta^\varepsilon + v_q^\varepsilon\right),$$

which increases with μ and decreases with k.

This result provides a benchmark for reforms that primarily change the profitability of invention. For example, reducing regulatory costs or increasing market size can increase the amount of invention without necessarily changing whether inventors pursue cost reduction or quality improvement. However, the prediction requires the reform to affect the returns

or costs of both types of R&D proportionally. A policy that disproportionately benefits one type of improvement can also change the direction of invention. Likewise, a fixed reduction in commercialization costs need not have the same effect on existing projects as a proportional reduction in the marginal cost of R&D. Whether reforms that change overall profitability leave the direction of invention unchanged is therefore an empirical question.

The model also predicts that increasing cost exposure can raise total R&D, since it increases the return to cost-reducing effort. However, actual payment reforms may also change market size, profitability, and other costs of innovation. Consequently, the model does not imply that reforms increasing providers' cost exposure must increase the total amount of invention.

Exposure Within a Reform

Payment reforms generally affect only some of the settings in which a device is used. A device used primarily in hospitals subject to a reform should therefore experience a larger change in demand for cost-reducing improvements than one used mainly in unaffected settings.

Let $w \in [0, 1]$ denote the share of a device's expected use that is subject to the reformed payment rule. Suppose the decision-maker bears a share θ_0 of the relevant cost under the original payment system and θ_1 under the new system. Holding the distribution of use and other valuations fixed, the average cost exposure relevant to the manufacturer is

$$\bar{\theta} = (1 - w)\theta_0 + w\theta_1.$$

Substituting this into the expression for the direction of invention gives

$$\frac{\partial s^\star}{\partial w} = \frac{(\theta_1 - \theta_0)v_q}{(\bar{\theta} + v_q)^2}.$$

When the reform increases cost exposure ($\theta_1 > \theta_0$), this derivative is positive. Thus, the model predicts larger shifts toward cost reduction among technologies used more intensively in affected settings.

This prediction motivates the comparisons between affected and unaffected technologies in Section 5. It does not imply that the magnitude of the response is proportional to the dollar value of the payment unit. The relevant incentive is how much the decision-maker benefits from an invention that reduces costs, which depends on the payment rule and the costs the invention can affect.

Several Margins of Cost Reduction

The baseline model treats cost reduction as a single objective. However, payment reforms differ in which costs they place on the decision-maker. In particular, a technology can reduce the costs of delivering care or the costs of producing the device itself. These are economically distinct improvements.

I therefore extend the model to three margins of invention: reducing episode costs (a_{epi}), reducing device production costs (a_{prod}), and improving quality (a_q). As before, the manu-

facturer faces the same quadratic cost of effort on each margin. Let the returns to the three objectives, before the common scale factor μ, be r_{epi}, r_{prod}, and r_q.

For episode-cost reduction, the mechanism is unchanged:

$$r_{\text{epi}} = \theta_{\text{epi}},$$

where θ_{epi} denotes the provider's exposure to the costs of delivering care. A technology that prevents complications, shortens hospital stays, or reduces follow-up care becomes more valuable when the provider bears more of those costs.

Production-cost reduction differs because it directly lowers the manufacturer's costs rather than the buyer's costs of using the device. A manufacturer benefits from reducing production costs on the units it already sells, even if the buyer does not observe the improvement and the selling price remains unchanged. This provides a return to production-cost invention regardless of the payment system.

However, lower production costs can also allow the manufacturer to offer a cheaper device or submit a lower bid for a purchasing contract. The value of this opportunity depends on whether buyers bear the device's purchase price. When hospitals pay for devices from a fixed reimbursement, lower prices can help manufacturers win sales. When hospitals receive separate reimbursement for the device, their purchasing decisions may be less sensitive to its price.

I summarize these two returns as

$$r_{\text{prod}} = \underbrace{r_0}_{\substack{\text{Savings on} \\ \text{existing sales}}} + \underbrace{\sigma\theta_{\text{dev}}}_{\substack{\text{Additional return from} \\ \text{price-sensitive sales}}}, \tag{3}$$

where $r_0 > 0$ is the direct return from reducing manufacturing costs, $\theta_{\text{dev}} \in [0, 1]$ is the buyer's exposure to the device's purchase price, and $\sigma \geq 0$ measures the additional return to lower-cost devices when buyers bear that price. This is a reduced-form representation of purchasing incentives, rather than a model of the bidding process itself.[57]

For quality improvement, I retain

$$r_q = v_q.$$

The manufacturer now solves

$$\max_{a_{\text{epi}}, a_{\text{prod}}, a_q \geq 0} \quad \mu \left(r_{\text{epi}} a_{\text{epi}} + r_{\text{prod}} a_{\text{prod}} + r_q a_q \right) - \frac{k}{2} \left(a_{\text{epi}}^2 + a_{\text{prod}}^2 + a_q^2 \right).$$

[57]The second term captures an increase in expected sales or profits from being able to offer a lower-priced device. It need not imply that all of the production-cost saving is passed through to buyers. The first term remains positive even when the buyer is fully reimbursed for the device. For simplicity, I assume both terms scale with the common market-return parameter μ. This is appropriate, for example, when μ changes with the number of units sold. It need not hold when a policy changes bargaining power or taxes the revenue from device sales, since those changes may affect the return to reducing manufacturing costs differently from the return to increasing buyer willingness to pay. In those cases, Proposition 2 need not apply to all three margins.

The optimal effort on each margin is

$$a_j^\star = \frac{\mu r_j}{k},$$

and its share of total R&D is

$$s_j^\star = \frac{r_j}{R}, \qquad R = r_{\text{epi}} + r_{\text{prod}} + r_q.$$

Thus, an increase in the return to one type of improvement raises its share of R&D while reducing the shares allocated to other improvements, holding their returns fixed.

Proposition 3 (Changes in relative returns redirect invention)**.** *For any margin j with a positive return, an increase in r_j, holding all other returns fixed, increases its share of R&D:*

$$\frac{\partial s_j^\star}{\partial r_j} = \frac{R - r_j}{R^2} > 0.$$

For every other margin $i \neq j$ with $r_i > 0$,

$$\frac{\partial s_i^\star}{\partial r_j} = -\frac{r_i}{R^2} < 0.$$

The distinction between absolute effort and its share is important. In this model, increasing the return to episode-cost reduction increases effort directed toward that objective without reducing the amount of effort devoted to quality or production-cost reduction. Their shares decline because episode-cost effort increases. The empirical patent scores are not literal R&D shares, however, so the model does not require their estimated changes to add up to zero.

The proposition gives clear predictions when a reform primarily changes one cost margin. If a reform increases θ_{epi} while leaving the return to production-cost reduction unchanged, invention shifts toward reducing episode costs. Conversely, when separate device reimbursement expires, the provider becomes more exposed to the device's purchase price. If $\sigma > 0$, this raises r_{prod} and shifts invention toward reducing production costs, while reducing the relative share of episode-cost invention.

Some reforms may change both cost exposures. For example, a bundled payment can increase a provider's exposure to the costs of delivering care and to device purchase prices. In that case, the direction of invention depends on which return increases more relative to the other returns. Let $\Delta r_{\text{epi}} > 0$ and $\Delta r_{\text{prod}} \geq 0$ denote the changes. Holding r_q fixed, the share of production-cost invention declines if and only if

$$\Delta r_{\text{epi}}\, r_{\text{prod}} > \Delta r_{\text{prod}} \left(r_{\text{epi}} + r_q\right).$$

Under this condition, the increase in the return to episode-cost reduction is sufficiently large that the episode-cost share rises while the production-cost share falls. The condition is not guaranteed simply because device purchases account for a small share of episode spending. It depends on how the reform changes the returns to developing each type of improvement.

For illustration, suppose the returns to episode-cost reduction, production-cost reduction, and quality improvement are initially 1, 0.5, and 2. The corresponding shares of R&D effort are 29, 14, and 57 percent. If a reform raises the return to episode-cost reduction from 1 to 1.5, holding the other returns fixed, these shares become 37.5, 12.5, and 50 percent. The episode-cost share therefore increases by approximately nine percentage points, while the shares of the other two objectives decline. This example illustrates the model's mechanism, not the quantitative interpretation of the patent-score estimates.

The different cost margins also have implications for who benefits from cost-reducing technologies. Under a fixed payment, a provider generally benefits directly from reducing the costs of an episode of care. By contrast, a manufacturer initially benefits from reducing its own production costs. How much of that saving reaches buyers depends on device prices, purchasing decisions, and competition. Lower production costs may lead to cheaper models or bids, while patent expiration can create additional opportunities for competitors to offer lower-priced devices. Section 7 examines these relationships empirically.

Who Redirects Invention

The model describes how the returns to different improvements affect the direction of R&D within a device category. It does not specify which manufacturers undertake that R&D. A change in the relative returns to cost reduction could lead existing firms to redirect their own research, or it could change which firms find it profitable to invent in the affected technologies.

The model therefore predicts the direction of the response, but not the relative contribution of within-firm changes and reallocation across firms. Section 6.3 examines these margins separately and finds evidence of both.

Related Theory

The model builds on the prediction that health care payment systems can shape the direction of medical innovation. Weisbrod (1991) argues that insurance and payment incentives may encourage technologies that expand medical capabilities without reducing costs. Goddeeris (1984) and Baumgardner (1991) develop related arguments in which insurance changes consumers' valuations of medical technologies.

The residual-claimancy parameter is closely related to models of provider payment, in which prospective payment makes providers responsible for more of the marginal costs of delivering care (Ellis and McGuire, 1986; Ma, 1994). Similar incentives arise in models of regulation and procurement (Laffont and Tirole, 1993). In those settings, the provider or contractor typically chooses its own cost-reducing effort. Here, the provider bears the cost, while an upstream manufacturer chooses R&D in response to the provider's willingness to pay.

The allocation of effort across improvements also relates to multitask models, in which the relative returns to different activities affect how effort is allocated (Holmstrom and Milgrom, 1991). Finally, the model relates to theories of directed technological change (Acemoglu, 2002; Acemoglu and Linn, 2004). Those models emphasize differences in market size across types of innovation. Here, I isolate a different mechanism: payment systems change the

relative value of cost reduction and quality improvement within a market. Market size can still affect direction if it changes the relative returns to these improvements, but a proportional increase in the returns to both does not.

Proofs

Proposition 1. The objective is strictly concave for $k > 0$. The first-order conditions give

$$a_c^\star = \frac{\mu\theta}{k}, \qquad a_q^\star = \frac{\mu v_q}{k}.$$

Since $v_q > 0$, the direction of invention is well defined, including when $\theta = 0$. Thus,

$$s^\star = \frac{\theta}{\theta + v_q}, \qquad \frac{\partial s^\star}{\partial \theta} = \frac{v_q}{(\theta + v_q)^2} > 0.$$

Proposition 2. With the common constant-elasticity cost function

$$C_j(a_j) = \frac{k}{1 + 1/\varepsilon} a_j^{1+1/\varepsilon},$$

the first-order condition for margin j is

$$\mu r_j = k a_j^{1/\varepsilon}.$$

Therefore,

$$a_j^\star = \left(\frac{\mu r_j}{k}\right)^\varepsilon.$$

Substituting $r_c = \theta$ and $r_q = v_q$ gives

$$s^\star = \frac{(\mu\theta/k)^\varepsilon}{(\mu\theta/k)^\varepsilon + (\mu v_q/k)^\varepsilon} = \frac{\theta^\varepsilon}{\theta^\varepsilon + v_q^\varepsilon}.$$

The common factors μ and k cancel from the share. Total effort is

$$a_c^\star + a_q^\star = \left(\frac{\mu}{k}\right)^\varepsilon \left(\theta^\varepsilon + v_q^\varepsilon\right),$$

which increases with μ and decreases with k.

Proposition 3. With $R = \sum_\ell r_\ell$ and $s_j = r_j/R$,

$$\frac{\partial s_j}{\partial r_j} = \frac{R - r_j}{R^2} > 0$$

provided at least one other margin has a positive return. For $i \neq j$ with $r_i > 0$,

$$\frac{\partial s_i}{\partial r_j} = -\frac{r_i}{R^2} < 0.$$

For a reform that changes both episode-cost and production-cost returns, the change in the episode-cost share is

$$\Delta s_{\text{epi}} = \frac{\Delta r_{\text{epi}}(r_{\text{prod}} + r_q) - r_{\text{epi}}\Delta r_{\text{prod}}}{R(R + \Delta r_{\text{epi}} + \Delta r_{\text{prod}})},$$

while the change in the production-cost share is

$$\Delta s_{\text{prod}} = \frac{\Delta r_{\text{prod}}(r_{\text{epi}} + r_q) - r_{\text{prod}}\Delta r_{\text{epi}}}{R(R + \Delta r_{\text{epi}} + \Delta r_{\text{prod}})}.$$

Thus, the production-cost share falls precisely when

$$r_{\text{prod}}\Delta r_{\text{epi}} > (r_{\text{epi}} + r_q)\Delta r_{\text{prod}}.$$

Since $r_q > 0$, this condition also implies an increase in the episode-cost share, completing the comparative-static result.

C Data Construction and Measure Validation

This appendix describes the construction and validation of the patent measures and the links between patents and the technologies affected by each reform. I first describe the patent sample and the classification of invention objectives. I then validate the measures against independent human coding and examine their sensitivity to alternative classifiers and labeling instructions. Finally, I describe the patent-to-payment crosswalks, the construction of treatment exposure, and the separate measure used for drug patents.

Patent Sample and Technology Classification

I identify medical device patents using the Cooperative Patent Classification (CPC), which organizes patents into technology classes, subclasses, main groups, and subgroups. Medical devices are concentrated in class A61, including diagnosis and surgery (A61B), prostheses and dressings (A61F), patient transport and accommodation (A61G), and devices for introducing substances into the body (A61M). A CPC main group is a narrower technology category within a subclass; for example, A61M25 includes catheters. I use main groups when exposure to a reform is defined by technology rather than through a patent-to-procedure crosswalk. Table A.2 lists the main groups and their official descriptions for each design.

I exclude patents whose primary CPC group covers consumer products or technologies not used primarily for human medical care. These include cosmetics and toiletries, absorbent articles such as diapers and sanitary products, massage and spa products, infant feeding bottles and comforters, veterinary devices, goggles and eye masks, ear protection, sexual aids, biometric identification, and male contraceptives. These exclusions remove 46,635 of 461,440 patents, or approximately 10 percent. I fixed the exclusions before estimation and apply them consistently across the supply-side designs.[58]

[58]Funeral equipment, including coffins, coffin hoists, and hearses (A61G17, A61G19, and A61G21), is

I also examine a stricter definition that excludes additional consumer and wellness technologies, including anti-snoring and nasal devices, nail and stammering aids, and several massage, bathing, and similar products. In the post-acute design, this specification also treats continence and ostomy supplies as a separate home-health category. The post-acute estimate increases from 1.11 to 1.29 percentage points, while no other estimate changes by more than 0.05 percentage points (Table A.12).

Constructing the Patent Measures

I measure the direction of invention using each patent's title, abstract, and brief summary of invention. The classifier assigns separate scores to six objectives: reducing device production costs, improving quality, preventing complications, shortening hospital stays, moving care to less costly settings, and reducing clinical labor. The scores range from zero to one and measure how strongly the patent text describes each objective. They are not calibrated probabilities that the invention belongs to a particular category, and a patent can receive high scores on several objectives.

I construct the measures in two steps. First, I use a language model to label a sample of patents according to the economic problems their inventions address. Second, I train a logistic regression classifier on those labels, representing patent text using term frequency–inverse document frequency (TF-IDF) weights. TF-IDF gives less weight to words that appear frequently across patents and more weight to words that distinguish individual documents. The resulting classifier scores the full patent sample without requiring a language-model evaluation of every patent.

The training sample contains 5,460 patents, drawn randomly within cells defined by grant period and CPC subclass, rather than by their text or invention objectives. I use gpt-5-mini through the Responses API with low reasoning effort to generate the labels. The model reads the same patent fields as the classifier. The prompt asks it to distinguish inventions that achieve a given clinical outcome at lower cost from those that expand medical capabilities or improve quality. It also instructs the model to identify the economic problem addressed by the invention rather than infer its purpose from the device category.

The labeling prompt is reproduced verbatim below. Importantly, it does not mention Medicare, the payment reforms, or the results being tested.

```
You are a medical-technology economist classifying what ECONOMIC PROBLEM a
  patented invention is designed to solve. Your job: separate
  episode-cost-reducing inventions from quality-expanding ones. Judge the
  PROBLEM SOLVED, not the device category.

PATENT TEXT:
{text}

Rate 0-1 how strongly the invention's PRIMARY PURPOSE is each channel (it can
  score on several):

EPISODE-COST-REDUCING:
```

excluded from the 1983 and post-acute designs and does not enter the BPCI or competitive-bidding samples. In the remaining designs, these groups account for at most 0.02 percent of linked patents.

```
- los_reducing: shortens hospital stay / speeds recovery / faster healing /
  earlier discharge.
  ('minimally invasive' counts here only if recovery/discharge benefit is
    implied.)
- site_shifting: enables a LOWER-acuity setting than before - outpatient /
  ambulatory / same-day /
  office / home / self-administered / point-of-care - avoiding hospitalization.
- complication_reducing: reduces complications / infections / adverse events /
  readmissions /
  device failures (the AIM is less harm or rework, not merely being
    'effective').
- labor_saving: reduces clinician time / staff count / required skill / steps
  / procedure duration.

CONTRAST (NOT episode-cost-reducing):
- quality_expanding: primary aim is NEW capability / higher efficacy /
  accuracy / performance /
  treating a previously-untreatable condition - typically at equal or HIGHER
    cost. The DEFAULT
  for most medical inventions.
- production_cost_reducing: cheaper to MANUFACTURE (fewer/cheaper parts,
  simpler process). Lowers
  the firm's production cost, NOT the episode cost.

CRITICAL: 'works better / more effective / more accurate / higher resolution'
  = quality_expanding,
NOT cost-saving - UNLESS it explicitly achieves the SAME clinical outcome with
  LESS resource
(shorter stay, fewer complications, outpatient, less labor). A device can be
  both; use graded scores.

Then: primary_channel = the SINGLE dominant economic purpose (or 'none' if
  purely incremental);
evidence_span = the verbatim phrase (<=200 chars) most supporting
  primary_channel (or '');
confidence 0-1; rationale = one sentence. Return ONLY the JSON object.
```

The model returns six scores, a primary invention objective, a supporting passage from the patent, and a short explanation. I train the classifier separately for each channel using the language-model labels, treating ratings of at least 0.5 as positive. The main production-cost and quality measures use their corresponding scores. The episode-cost measure averages the scores for shorter hospital stays, fewer complications, and treatment in less costly settings. For the 1983 inpatient reform, I use only the first two channels, since prospective payment initially made hospitals responsible for costs during the inpatient stay rather than the costs of subsequent care. I examine labor saving separately because I cannot validate it against a corresponding outcome in the available claims data.

Table A.4 reports classification performance against held-out language-model labels. I measure performance using the area under the receiver operating characteristic curve (AUC), which is the probability that a randomly selected positive patent receives a higher score than a randomly selected negative patent. An AUC of 0.5 indicates no discriminatory power, while an AUC of 1 indicates perfect separation.

Against the labeling model's held-out classifications, the production-cost measure achieves an AUC of 0.93, the episode-cost measure 0.92, and the quality measure 0.79. Among the episode-cost channels, complication reduction has the lowest AUC (0.81). These results show how well the classifier reproduces labels generated by the language model. They do not independently establish whether those labels correctly describe the inventions, which I examine below using human coding.

The measures also differ from conventional patent classifications. Among 350,542 patents with a process-versus-product classification, the production-cost and episode-cost measures have little agreement with the Bena–Simintzi distinction (Cohen's $\kappa = -0.01$ and 0.04, respectively). Thus, identifying inventions that reduce costs is not equivalent to identifying process innovations. Within technology class and year, 84 percent of the variance in the production-cost score occurs across individual patents (Table A.13). The measure therefore contains substantial variation within technology fields rather than simply assigning different scores to different types of devices.

Table A.14 reports the eight terms receiving the greatest weight in each classifier. Production-cost invention is associated with words such as "cost," "inexpensive," and "manufacturing." The stay-reducing measure emphasizes "healing," "incision," and "wound," while site shifting emphasizes "portable," "ambulatory," and "home." The quality measure has a less distinctive vocabulary, with terms referring to particular medical applications and capabilities. These patterns are consistent with the intended classifications, although the term weights alone do not establish their accuracy.

Validation Against Independent Human Coding

I validate the classifier using approximately 200 device patents independently rated by a research assistant, including 150 that I also coded. The sample was stratified by predicted production-cost scores, with high-scoring patents oversampled to provide sufficient examples of cost-reducing invention. For my own ratings, I did not observe the patent numbers, the research assistant's classifications, or the classifier's scores. Each coder rated the extent to which a patent targeted each invention objective, from zero to one. I classify a patent as targeting an objective when its rating is at least 0.5.

Against the research assistant's coding, the production-cost measure achieves an AUC of 0.82, compared with 0.68 for quality improvement and 0.66 for complication reduction. On the 150 patents rated by both coders, the production-cost measure achieves an AUC of 0.84 against their average ratings. The corresponding AUCs are 0.78 for quality and 0.76 for complications (Table A.15). Reweighting the sample to represent the population of device patents gives AUCs of 0.74, 0.79, and 0.74, respectively. The difference between the unweighted and weighted estimates matters because the validation sample deliberately oversamples patents with high classifier scores.[59]

Agreement between the human coders is itself imperfect. For production cost, their binary classifications have Cohen's $\kappa = 0.44$, while the classifier agrees with the two coders

[59]For production cost, the AUC ranges from 0.84 to 0.88 depending on whether I use either coder alone or their combined ratings. The corresponding ranges are 0.51–0.78 for quality improvement and 0.68–0.77 for complication reduction. Two of the 150 double-coded patents appear in the classifier's training sample; excluding them changes no AUC by more than 0.005.

at $\kappa = 0.45$ and 0.66. Cohen's κ measures agreement beyond chance, with zero indicating chance-level agreement. The lower agreement on quality improvement and complication reduction reflects, in part, differences in how broadly the coders interpret these objectives.

I also examine which reader disagrees when the classifier and human coders give different classifications. Among the 48 patents with disagreement over production-cost reduction, the classifier differs from both coders on 11, compared with 24 for the first coder and 13 for the second. When the first coder disagrees with the classifier, the second agrees with the classifier on 24 of 35 patents. For complication reduction, the classifier differs from both coders on 18 of 94 disputed patents. For quality improvement, it differs from both on 63 of 99. The larger disagreement over quality reflects differences in classification thresholds: both human coders classify approximately 80 percent of patents as quality-improving, compared with fewer than 40 percent of patents receiving classifier scores above 0.5. Nevertheless, the classifier ranks quality-improving patents similarly to the coders, as reflected in the AUC of 0.78.

Some disagreements arise because the human classifications do not follow the same economic distinction as the measure. For example, one coder classifies a dental prosthesis described as "economical to produce" and mask headgear using a "relatively low cost manufacturing technique" as not reducing production costs. The classifier assigns them production-cost scores of 0.80 and 0.71. Conversely, the coder classifies a transgastric hernia-repair system as production-cost-reducing because it makes treatment less expensive. However, the savings described in the patent come from shorter hospital stays and fewer complications rather than cheaper manufacturing. The classifier assigns a complication-reduction score of 0.88 and a production-cost score of 0.07.

Other disagreements are less clear. A patent for an additively manufactured denture base describes lower material and labor requirements but also emphasizes improvements in fit. Similarly, a heater printed onto a medical fluid bag uses inexpensive electronics to make a new application feasible. In both cases, the coder emphasizes quality improvement, while the classifier assigns relatively high production-cost scores (0.82 and 0.75). These examples illustrate that some inventions pursue several objectives and that the distinction between lowering costs and enabling new capabilities is not always clear from patent text.

I further compare the classifier with labels from four language models (gpt-5-mini, gpt-5, Claude, and gpt-5.5) on a separate sample of 643 patents. Pairwise agreement among the models ranges from $\kappa = 0.40$ to 0.77, and from 0.61 to 0.77 among gpt-5-mini, gpt-5, and Claude. On the 202 patents with human coding, the classifier's agreement with the human coder on production-cost reduction is similar to that of a frontier language model ($\kappa = 0.39$ versus 0.38; Table A.15). The models generally agree with one another more than they agree with the human coder. These comparisons provide an additional benchmark, but agreement among language models is not an independent measure of accuracy.

Robustness of the Patent Measures

I examine three concerns: sensitivity to the classifier, sensitivity to the labeling instructions, and changes in classification error around the reforms.

Alternative classifiers. I train three alternative classifiers using the same language-model

labels and patent text: a general-purpose GTE embedding, the patent-specific PaECTER embedding, and a DistilBERT sentence encoder. The baseline TF-IDF classifier has the highest performance against the held-out training labels across the main channels, so I use it for the analysis and treat the alternatives as robustness checks (Table A.16).

The principal results are not specific to TF-IDF. All four classifiers produce positive and statistically significant estimates for the 1983 and BPCI reforms. Across the four measures, the 1983 estimate ranges from 2.2 to 4.5 percentage points and the BPCI estimate from 1.2 to 2.1. The estimates for the other reforms generally retain their signs, although the production-cost response to pass-through expiration is statistically significant only under the baseline classifier. The episode-cost response at pass-through expiration is zero under GTE, the least accurate classifier, rather than negative.

The classifiers also perform similarly on an independent set of 146 patents classified from their full text by Claude Opus. For TF-IDF, PaECTER, and GTE, the AUCs are 0.89, 0.88, and 0.85 for episode-cost reduction, and 0.90, 0.90, and 0.86 for length-of-stay reduction. Complication reduction remains less accurately measured, with AUCs of 0.69, 0.67, and 0.68. Against the research assistant's coding, TF-IDF is more accurate on production costs (AUC 0.82 versus 0.72–0.76 for the alternatives).[60]

Labeling instructions. A potential concern is that language-model labels could depend on how the research question is described. The baseline prompt deliberately avoids mentioning Medicare or any payment reform. To test whether this matters, I generate a separate set of labels using a prompt that explicitly describes the study as examining payment incentives, identifies Medicare's 1983 DRG reform, and labels the episode-cost channels as "what DRG rewards." The definitions of the invention channels are otherwise unchanged.

Labels from the two prompts are correlated at 0.83 for production cost and 0.82 for episode-cost reduction, compared with correlations of 0.88 and 0.87 across repeated draws of the baseline prompt. Both prompts have similar agreement with human coding: production-cost AUCs of 0.82 and 0.84, and complication AUCs of 0.75 and 0.76.

However, explicitly mentioning the reform changes some estimated effects. The 1983 estimate increases from 3.31 to 4.39 percentage points (difference 1.08, s.e. 0.35), while the episode-cost estimate following pass-through expiration becomes more negative, changing from -0.91 to -1.32 percentage points (difference -0.40, s.e. 0.10). Most other estimates remain similar: BPCI changes from 1.58 to 1.45 percentage points, the post-acute estimate from 1.11 to 1.25, and production-cost invention after pass-through expiration from 1.02 to 1.10.

Thus, naming a reform in the prompt can change the estimated response in the direction suggested by the reform's incentives, even when the labels have similar agreement with human coding. For this reason, I use the reform-independent prompt reproduced above for every device patent.

Changes in measurement error over time. A classifier that consistently ranks patents according to their objectives could still produce biased estimates if its errors change dif-

[60]These comparisons exclude patents appearing in the classifiers' training data, leaving 199 human-coded patents for production cost and 146 independently model-coded patents for complications and length of stay. Against the independent model labels, DistilBERT performs slightly better than TF-IDF on complications (AUC 0.72 versus 0.69), while the two are similarly accurate on length of stay (0.90).

ferently in affected and comparison technologies after a reform. I examine this possibility using the 5,460 training patents, for which I observe both the language-model labels and out-of-sample classifier scores.

For each reform, I estimate the main specification using the part of the classifier score not explained by its training label. None of the estimated treatment effects is statistically significant. For the 1983 reform over 1991–1995, the estimate is +0.4 percentage points (s.e. 3.5); for BPCI, −4.8 (s.e. 7.2); and for the post-acute reforms, +0.7 (s.e. 1.6). These estimates are imprecise: the confidence intervals cannot rule out differential changes in error several times larger than the estimated reform effects. Moreover, because the labels are generated by the model used to train the classifier, this exercise tests errors introduced by the TF-IDF classifier, not errors in the underlying language-model labels.

I also examine whether the classifier's performance changes across patent cohorts. Discrimination against held-out labels is similar before and after 2005, although the average production-cost score declines over time, as does the fraction of patents labeled as production-cost-reducing by the language model. The empirical specifications include year fixed effects, which account for changes common to all technologies in a year. They do not rule out differential changes in patent language across affected and comparison technologies.

A separate check uses the language-model training labels directly, before applying the TF-IDF classifier. Among 991 training patents in technology fields included in the 1983 design, the fraction labeled as reducing complications or length of stay rises by 12.9 percentage points more in inpatient fields than in comparison fields between the pre-reform period and 1991–1995. This corresponds to approximately 3.7 percentage points on the classifier-score scale.[61] However, the shift in the training labels is already present during 1983–1990 and does not increase further after 1990–1991. This check supports the direction of the estimated response, but not the later increase in its magnitude. Because the labels come from the same language model used to train the classifier, it also does not independently validate their economic interpretation.

For comparison, I construct a simpler measure that counts keywords associated with complications and length of stay. This measure produces no detectable response after the 1983 reform. The estimated shift therefore depends on information extracted by the trained classifier rather than simple keyword frequencies.

Finally, the measures are largely distinct from conventional indicators of patent quality. Within technology class and issue year, correlations between the production-cost or episode-cost scores and measures of patent value, novelty, and citation impact are all below approximately 0.10 in absolute value (Table A.17). Applying the BPCI design to patent novelty produces a small negative estimate, but the outcome exhibits significant differential pre-reform trends. The probability of a breakthrough patent shows no significant change.[62]

[61] The sample contains 991 patents across 30 fields, including 126 patents filed before 1983. I estimate the main specification with field and filing-year fixed effects and standard errors clustered by field. The 1991–1995 estimate is +0.129 (s.e. 0.043, $p = 0.006$); the 1983–1990 estimate is +0.137 (s.e. 0.047, $p = 0.008$). A unit increase in the fraction of positive labels is associated with a 0.29 increase in the classifier score.

[62] I measure novelty as minus the log of backward patent similarity (Kelly et al., 2021). The BPCI estimate is −0.06 standard deviations ($p = 0.02$), but the pre-reform coefficients are jointly significant ($p < 0.001$). Within technology class and issue year, the estimate is −0.04 ($p = 0.17$). For the breakthrough indicator of Kelly et al. (2021), the estimate is −0.1 percentage points ($p = 0.89$), with no significant pre-reform trend

Patent-to-Payment Crosswalks

The empirical designs require linking patents to the medical procedures, devices, or drugs affected by each reform. These classifications are not recorded in patent data, so I construct crosswalks using text embeddings that measure the similarity between patent descriptions and the descriptions of billing codes or FDA drug applications.

The main concern is not simply whether every link is correct. Incorrect links may weaken the relationship between the treatment and the technologies it affects, but they could also bias the estimates if link accuracy differs systematically between treated and comparison technologies. I therefore examine both overall accuracy and differential accuracy.

Patent-to-HCPCS links. I evaluate the patent-to-HCPCS crosswalk using a stratified sample of 300 patent–billing-code links, selected across treated and comparison codes and three ranges of cosine similarity. The links are assessed without showing the coders their treatment status or similarity scores. I manually classify 72 links, while a language model classifies all 300. The two classifications have Cohen's $\kappa = 0.61$.

Table A.18 reports the results. The probability that a link refers to the same clinical domain increases sharply with similarity. In the lowest tier, approximately half of manually coded links and two-fifths of model-coded links are in the same clinical domain. In the highest tier, the corresponding share is approximately 90 percent. The relationship supports excluding links below the similarity threshold of 0.45 and restricting each patent to its closest matches, although some links above the threshold remain incorrect.

I next compare link accuracy between treated and comparison codes, controlling for similarity tier. Under the same-clinical-domain criterion, the treatment coefficient is $+0.06$ ($p = 0.59$) in the hand-coded sample and -0.01 ($p = 0.78$) in the model-coded sample. Under a stricter criterion requiring an exact code match, the estimates are $+0.14$ ($p = 0.21$) and -0.06 ($p = 0.21$). None is statistically significant. However, the samples can detect only relatively large differences in accuracy, approximately 0.29 for the hand-coded sample and 0.13 for the model-coded sample. The audit therefore provides no evidence of differential linking error, but does not rule out smaller differences.

Exposure to the 1983 reform. For the 1983 inpatient reform, I assign exposure using the medical setting in which each CPC main group's technologies were primarily used before the reform. I classify a group as inpatient if its devices were used mainly during an admitted hospital stay in 1976–1982, and as a comparison technology if its devices were used mainly outside inpatient care. The classification uses each group's official CPC description and a rule established before estimation. It does not use patent outcomes or language-model ratings.

The resulting sample contains 27 inpatient groups and 30 comparison groups. I exclude technologies whose setting is ambiguous, along with dental groups and non-device patents. Table A.2 lists the classifications. The estimated increase after 1990–1991 is not sensitive to the year used to separate the two post-reform periods. Using split years from 1987 to 1994, the later-period estimates range from 2.5 to 3.9 percentage points, each statistically significant at the 5 percent level.

($p = 0.60$).

I also construct a continuous measure of inpatient exposure using a language model. For each technology group, I ask gpt-5-mini to estimate the share of U.S. medical use that occurred during an inpatient hospital stay around 1980. The prompt provides the CPC code and description, includes three anchors for the rating, and explicitly asks the model to describe use at that time rather than current practice. It does not mention Medicare, prospective payment, or the reform. I obtain five ratings per group and average them; individual ratings have an average correlation of 0.92 across draws.[63] A one-standard-deviation increase in this exposure measure is associated with a 1.73-percentage-point larger increase in the episode-cost score during 1991–1995 (s.e. 0.57).

I compare both exposure measures with Medicare claims, although these data are available only several decades after the reform. Using the Medicare Physician/Supplier Procedure Summary for 2010–2012, I calculate the share of services billed in an inpatient hospital for each procedure code. I link these shares to patents and average them within CPC groups using patents filed during 1976–1982.

Among the 35 classified groups with a claims-based measure, the average inpatient share is 17 percent for groups classified as inpatient and 2 percent for comparison groups. At the procedure-code level, the corresponding shares are 54 and 8 percent. The continuous language-model measure correlates with the claims-based measure at 0.50, or 0.72 when weighted by patent counts. Using claims-based exposure in the regression gives a smaller and imprecise estimate (0.69, s.e. 0.62, compared with 0.97 on the same groups). These claims provide a useful check on the classification, but they do not directly measure technology use before 1983 and omit some hospital-supplied equipment. I therefore retain the original exposure measures.

BPCI exposure. For bundled payment, I link each patent to its five closest procedure codes and classify it as exposed if at least one code belongs to a BPCI episode. I use multiple links because a device may be used in several procedures, only some of which were covered by the reform. In the pass-through analysis, by contrast, I use the single closest billing code because each device category has its own start and expiration dates.

The results do not depend on the exact number of procedure links or the similarity threshold. Using only the closest link produces a BPCI estimate of 1.13 percentage points (s.e. 0.43). Across all 25 combinations of one to five links and cosine thresholds from 0.40 to 0.50, estimates range from 0.89 to 1.72 percentage points and are statistically significant at the 5 percent level. The threshold affects the sample size, which ranges from approximately 34,000 to 76,000 patents.

The response does not increase monotonically with the share of a patent's procedure links covered by BPCI. Among exposed patents, those with low, medium, and high covered shares (averages of 0.22, 0.40, and 0.76) have estimated increases of 1.62 (s.e. 0.53), 1.62 (s.e. 0.82), and 1.29 percentage points (s.e. 0.58), respectively. Thus, the estimates are robust to alternative exposure definitions, but the continuous measure does not provide evidence of a larger response among patents with more covered links.

[63] For twelve small groups, covering 254 patents, I used a shortened CPC title rather than the official description. Excluding them gives an estimate of 1.76 percentage points per standard deviation of exposure, compared with 1.73 when they are included. A separate prompt that mentions the 1983 reform and examples of procedures that subsequently moved to outpatient settings produces exposure ratings correlated at 0.93 with the baseline five-rating average.

The preferred comparison group consists of patents linked to major device-intensive surgical procedures not covered by BPCI. Patents linked only to supplies, equipment, drugs, vision, or dental codes are not included. This restriction makes the comparison technologies more similar to those used in bundled procedures.

Using all procedure-linked patents as comparisons gives an estimate of 1.29 percentage points (s.e. 0.44; pre-reform $p = 0.75$). Restricting comparisons to patents with a surgical link, patents whose closest link is surgical, procedures in the same body systems, or technology classes with an exposed patent before 2010 gives estimates from 1.43 to 1.62 percentage points, all significant at the 1 percent level. I also test whether the preferred comparison group appears to lose cost-reducing invention because firms shift R&D toward bundled technologies. The difference between the preferred surgical comparisons and other unexposed patents is similar among firms with and without exposed patents, providing no evidence of such displacement.

Selection into text-linked exposure. A further concern is that reforms may change how inventors describe their technologies, which could affect whether a patent is linked to a treated procedure or billing code. For example, a manufacturer developing a cost-reducing device after BPCI might describe it using language more closely related to bundled procedures. This could change measured treatment exposure even if the underlying technology were unchanged.

I test whether the share of patents linked to treated procedures increases after the reforms, using technology classes with established links before treatment. I find no significant increase in treated linking rates. I also test whether the relationship between a patent's cost-reduction score and its likelihood of being linked changes after treatment. The estimated changes are 0.09 for BPCI ($p = 0.27$) and -0.01 for pass-through expiration ($p = 0.83$).

Finally, I assign BPCI exposure at the technology-class level using only patents filed during 2005–2009, before the reform was announced. This avoids allowing changes in post-reform patent descriptions to determine which classes are treated. The estimated increase remains statistically significant ($p < 0.001$). Entry into the procedure-linked BPCI sample also does not change significantly after treatment. These checks reduce concern that the estimated response is driven by changes in which patents are linked to affected procedures, although they cannot rule out every form of differential linking error.

Pass-through comparison groups. The preferred pass-through specification compares patents linked to device categories receiving separate reimbursement with patents linked to codes that never received it. Restricting the comparison group to the 4,699 codes not yet billed in 2000 gives a production-cost estimate of 1.02 percentage points (s.e. 0.35), identical to the baseline. Restricting comparisons to HCPCS Level II device and supply codes gives 1.18 percentage points (s.e. 0.35).

Changes Within Technologies and Firms

An increase in the average cost-reduction score can occur because patents within a technology become more focused on reducing costs, or because invention shifts toward narrower technologies that already have higher scores. Appendix Table A.9 examines these possibilities using CPC subgroups, the finest technology classifications in the data.

With subgroup fixed effects, the BPCI estimate is 1.3 percentage points, compared with 1.6 in the baseline. Holding the pre-reform distribution of subgroups fixed gives a similar estimate. The pass-through production-cost estimate is 0.7 percentage points within subgroups, approximately three-quarters of the baseline. For the 1983 and post-acute reforms, the within-subgroup estimates are smaller: 1.3 of 3.3 percentage points and 0.5 of 1.1 percentage points, respectively. Holding subgroup composition fixed gives similar results, suggesting that shifts toward neighboring technologies also contribute to the estimated changes.[64]

These results distinguish invention within narrowly defined technologies from changes in the mix of technologies being developed. However, CPC subgroups are themselves assigned according to patent content. A reform that changes the kinds of inventions developed may therefore change their subgroup classifications. Holding subgroups fixed can consequently remove part of the response rather than isolate measurement error.

I also restrict the sample to company-assigned patents, firms that patented before each reform, and firms that continue patenting after treatment. Across these specifications, the estimated effects range from approximately three-quarters to 1.3 times their baseline magnitudes. Thus, changes in the composition of firms do not account for the entire response. Section 6.3 examines within-firm changes and reallocation across firms in greater detail.

Drug Patent Measure and Crosswalk

The prescription-to-OTC analysis requires a separate measure because the device classifier was trained on medical-device patents and its episode-cost channels are not well suited to drugs purchased directly by patients.

I follow the same classification procedure wherever possible. I use gpt-5-mini with low reasoning effort to label drug patents based on their titles, abstracts, and summaries of invention. I then train separate TF-IDF logistic classifiers on the resulting labels, using a threshold of 0.5. The training sample contains 7,999 A61K drug patents granted between 1976 and 2024, and I apply the classifiers to 159,361 patents.

The drug classifier measures three objectives: reducing production costs, reducing patients' costs of using drugs, and improving quality. The patient-use-cost measure replaces the device-specific episode-cost channels. It captures inventions such as less frequent dosing, easier self-administration, and reduced monitoring requirements. As with the device classifier, the labeling instructions distinguish accomplishing the same therapeutic objective at lower cost from expanding therapeutic capabilities.

The drug labeling prompt is reproduced verbatim below.

```
You are labeling a PHARMACEUTICAL/DRUG patent for the DIRECTION of its
  innovation.
Read the TITLE, ABSTRACT, and SUMMARY OF INVENTION. Score each dimension
  0.0-1.0 = probability the
```

[64]The within-subgroup estimates have $p = 0.054$ for 1983, $p = 0.16$ for the post-acute reforms, $p < 0.001$ for BPCI, and $p = 0.014$ for pass-through production-cost invention. The pass-through episode-cost estimate is -0.53 percentage points ($p = 0.094$). Holding the pre-reform subgroup distribution fixed gives estimates of 1.51 percentage points ($p = 0.080$) for 1983, 0.61 ($p = 0.10$) for the post-acute reforms, and 1.33 ($p < 0.001$) for BPCI. I do not calculate this specification for pass-through because 58 percent of post-reform patents occur in codes without pre-reform patents.

```
patent's PRIMARY economic objective is that direction (a patent may score on
  more than one):

- production_cost_reducing: lowers the RESOURCE COST OF MAKING the drug at a
  GIVEN therapeutic quality
  - cheaper/simpler synthesis or process, higher yield, fewer/cheaper reagents
    or steps, improved
  manufacturability, crystalline-form/formulation change that cuts production
    cost, generic-enabling process.
- patient_use_cost_reducing: lowers the PATIENT'S cost/burden of USING the
  drug at GIVEN efficacy
  - lower or less-frequent dose, self-administration, easier delivery/adherence,
  reduced monitoring or clinic visits.
- quality_expanding: raises efficacy/performance/safety or adds a NEW
  indication or mechanism
  - better therapeutic effect, novel compound/MOA, expanded use, improved
    safety/side-effect profile.

Constant-quality rule: production_cost and patient_use_cost = doing the SAME
  therapeutic job more cheaply;
quality_expanding = doing MORE. Return STRICT JSON only:
{"production_cost_reducing": <float>, "patient_use_cost_reducing": <float>,
  "quality_expanding": <float>,
 "rationale": "<=25 words"}
```

In five-fold cross-validation against the language-model labels, the drug classifier achieves AUCs of 0.90 for production-cost reduction, 0.87 for patient-use-cost reduction, and 0.86 for quality improvement. As with the device classifier, these figures measure agreement with the training labels rather than independent accuracy. Unlike the device classifier, the drug measure has not been evaluated against human-coded patents.

I provide an external validation using FDA approval classifications. Among 4,240 scored patents listed in the Orange Book, patents protecting new dosage forms or formulations have an average patient-use-cost score of 0.65, compared with 0.36 for patents associated with new molecular entities. The difference is 1.4 standard deviations of the measure within filing year. Among oral products, patents protecting modified-release formulations score 1.8 standard deviations higher than those protecting immediate-release tablets or capsules. These patents also score lower on quality improvement, while their production-cost scores do not differ (Figure 6, panel (a)). Because FDA approval classifications are assigned independently of the patent text, these comparisons provide a separate check that the measure captures differences in the objectives of drug invention.

I link drug patents to FDA new-drug applications using the same text-embedding approach as the device crosswalk. I retain links with cosine similarity of at least 0.45 and each patent's five closest applications. Applications are linked to their active ingredients, and the treated unit is the ingredient form that switches to OTC status. If a patent is linked to several molecules within a matched event stack, I assign it to the most similar molecule.

I assess the drug crosswalk against manufacturers' patent listings in FDA's Orange Book. Among 5,605 listed patents that also receive crosswalk links, the application ranked first by text similarity has the same active ingredients as the listed application for 55 percent of patents above the 0.45 threshold. Accuracy increases with similarity, from 20 percent for

scores between 0.45 and 0.50 to 69 percent above 0.50. It is similar for patents linked to switched and never-switched molecules (57 and 55 percent).

The comparison is demanding because patents may cover related compounds or multiple applications, while the test requires matching the active ingredients of a particular listed application. Moreover, the available Orange Book listings disproportionately represent recent patents, with a median filing year of 2017. Nevertheless, the 55 percent match rate indicates substantial noise in the patent-to-drug links. The similar accuracy for switched and comparison molecules reduces concern about differential measurement error, but does not eliminate it. I therefore interpret the OTC results as suggestive evidence that the payment mechanism extends to patients rather than as a definitive test of the mechanism.

D Supplemental Welfare and Incidence Evidence

This appendix provides additional evidence on the relationship between cost-reducing invention and device prices, the robustness of the patent-expiration results, and how production-cost savings may be divided between manufacturers and buyers. I also describe the classification of device patents by payment setting used in Section 7.4.

Production-Cost Invention and Device Prices

I examine whether device prices fall more rapidly in markets with greater production-cost invention. I use ECRI hospital procurement data from 2012–2022, covering approximately 3.8 million catalog-item-year price observations. Because devices differ in their characteristics, I construct a quality-adjusted price index using descriptions of each product's type, materials, and function. Specifically, I regress log transaction prices on components of text embeddings of these descriptions and facility fixed effects. The resulting index adjusts for differences in the product characteristics captured by the descriptions, although it may not account for all differences in quality.

I measure each device category's exposure to production-cost invention using a text-embedding crosswalk between patents and FDA product codes. This crosswalk links approximately 785,000 A61 patents to 5,489 FDA product codes, using similarity between patent text and each code's definition. For each product code, I calculate the average production-cost score of patents filed by 2012, before the period over which I measure price changes. I use this predetermined measure to examine whether prices subsequently change differently across categories with more or less production-cost invention.[65]

Figure A.4 shows that quality-adjusted prices decline more rapidly in categories with higher production-cost scores. The price difference associated with a one-standard-deviation higher score grows from approximately −12 percent in 2013 to −33 percent in 2022. The difference is statistically significant in each year. With product-code fixed effects and a control for changes associated with quality-improving invention, the estimated price trend

[65] I retain patent-to-product-code links with cosine similarity of at least 0.45. I use the embedding crosswalk rather than the keyword-based crosswalk from the FDA reclassification analysis, which was curated for treated device codes and links disproportionately to those categories. The price index uses MiniLM embeddings of product descriptions.

is 4.3 percent lower per year for each standard deviation of the production-cost score when I use 50 embedding components ($p < 0.001$).

The relationship remains negative with more detailed controls for product characteristics. Using 150 embedding components reduces the estimated annual difference to 2.9 percent. Excluding 2016, an anomalous year in the ECRI data, gives 2.7 percent; weighting patent links by cosine similarity gives 2.8 percent; and restricting the sample to the years before COVID gives 4.5 percent. These estimates suggest that the relationship is not driven entirely by observed differences in product characteristics. However, the embedding controls cannot rule out differences in unobserved quality or other market characteristics associated with production-cost invention.

I next examine whether the price difference reflects lower prices for individual products or changes in which products hospitals purchase. These mechanisms have different implications for the incidence of cost reduction. If manufacturers lower the prices of existing devices, buyers benefit directly from lower transaction prices. Alternatively, buyers may lower their average spending per device by purchasing cheaper products, even if the prices of individual products remain unchanged.

The ECRI data show substantial price stickiness. Among 1.94 million consecutive-year observations of the same catalog item purchased by the same hospital, prices change by less than 0.1 percent in 50.5 percent of cases. However, 41.4 percent change by more than 1 percent, with a median absolute change of 5.6 percent among these observations. The average period between price changes is approximately 2.4 years.

I also examine prices around increases in production-cost patenting within device categories. Figure A.9 compares continuing products in categories experiencing large increases in cost-reducing invention with products in categories without comparable increases in either cost-reducing or quality-improving invention. Prices of continuing products change by less than 1.5 percent in each of the first four years after the increase. Average changes during years 0–3 and 4–6 are small and statistically insignificant. Increases in quality-improving patenting, which frequently coincide with increases in production-cost patenting, are followed by similar price paths. Thus, I find little evidence that increases in cost-reducing patenting lead manufacturers to lower the prices of their existing products.

A separate comparison of annual price changes within the same catalog item and hospital finds a negative association with the product code's production-cost score (-1.8 percent per standard deviation, $p = 0.001$). However, this cross-sectional relationship does not establish that prices fall in response to new cost-reducing inventions, as the event-study estimates do not show a comparable change after increases in patenting.[66]

The larger difference appears in the products hospitals purchase. When each product receives equal weight, the annual quality-adjusted price difference associated with a one-standard-deviation higher production-cost score is -0.3 percent ($p = 0.65$). When products are weighted by the quantities hospitals purchase, the corresponding estimate is -4.3 percent. Restricting the sample to products already on the market gives an estimate of -3.9 percent. Thus, the lower average prices paid in categories with more production-cost invention arise

[66]The within-item comparison includes year fixed effects and clusters standard errors across 494 product codes. Using a keyword-based patent link gives an estimate of -3.5 percent. The event studies in Figure A.9 instead examine price changes around increases in cost-reducing patenting.

primarily because hospitals purchase relatively more of the cheaper available products, rather than because the prices of individual products decline faster.[67]

I also examine whether manufacturers with more production-cost invention reduce their own prices. Using 736 manufacturers, I find no statistically significant relationship between a manufacturer's production-cost score and the subsequent prices of its products. The estimate is −1.5 percent per standard deviation per year within manufacturers ($p = 0.60$) and approximately zero across manufacturers ($p = 0.99$).[68]

Together, these results show that hospitals pay lower average prices in device categories with more production-cost invention, largely through changes in their purchasing decisions. However, the estimates do not establish that cost-reducing inventions cause this reallocation. Hospitals facing greater budget pressure or different contracting arrangements may both purchase cheaper devices and operate in markets where manufacturers pursue more cost reduction. The absence of a significant own-price response is consistent with manufacturers retaining some production-cost savings as margin, but it does not identify how much of the savings they retain.

Patent-Expiration Robustness

Section 7 examines whether device prices decline when patents associated with production-cost invention expire. I identify 24 product codes with large expirations of production-cost patents and match each to three comparison codes using Mahalanobis distance on price levels, price changes, quantities, and product counts in the preceding year.

I exclude comparison codes with a large expiration of either production-cost or quality-improving patents within three years of the treated event. This restriction matters because the two kinds of expiration are highly correlated (0.88 across product-code-years). If I screen only for expirations of the same type, approximately 5 percent of the candidate controls for production-cost expirations and 9 percent for quality-improving expirations have a large expiration of the other type nearby. Including these codes could make treated and comparison categories more similar after the event, reducing the estimated difference.

The estimated price decline following production-cost patent expiration is negative across the alternative comparison groups. Using codes that never experience a large expiration of the same type gives an expiration-year estimate of −0.103 log points (s.e. 0.063). Also excluding codes with nearby expirations of the other type gives −0.099 (s.e. 0.064). Using all eligible codes and excluding only nearby expirations of the same type gives −0.108 (s.e. 0.057). Finally, an unmatched comparison with approximately 500 codes without large expirations of either type within six years gives −0.068 (s.e. 0.026).[69]

The estimates following quality-improving patent expiration are generally smaller in mag-

[67]The raw aggregate price index, before adjusting for product characteristics, falls by approximately 68 log points over the sample period. Approximately 59 percent of this decline reflects changes in the product mix that the hedonic adjustment removes.

[68]The within-manufacturer specification uses catalog-item-by-hospital-by-year observations and includes manufacturer, facility, and year fixed effects. The across-manufacturer specification includes the manufacturer's production-cost score and year fixed effects.

[69]The corresponding estimates for quality-improving patent expirations are +0.012 (s.e. 0.071), −0.026 (s.e. 0.062), −0.014 (s.e. 0.052), and −0.060 (s.e. 0.028), respectively.

nitude, but I cannot statistically distinguish them from the production-cost estimates. The p-values for the differences range from 0.12 to 0.78. A specification including both types of expiration as continuous measures also cannot separately identify their effects, reflecting the substantial overlap between the two.

Another limitation is the measurement of patent expiration dates. I use the nominal expiration date, approximately 20 years after filing, but patents may expire earlier because maintenance fees are not paid. Actual terms can also differ for other legal reasons. Errors in the timing of expiration may spread any price response across years and reduce the measured change at the nominal date. However, the direction of any resulting bias is not guaranteed. The estimates also rely on linking patents to product codes, rather than observing which patents protect each purchased product. I therefore interpret the price decline as suggestive evidence that expiration of production-cost patents may increase access to cheaper devices, rather than a precise estimate of the price effect of patent expiration.

I next examine how the average price decline occurs. Figure A.7 decomposes changes in the quantity-weighted average price of each product code into four components: changes in the prices of catalog items purchased in both periods; changes in purchase shares among those continuing items; purchases of entering items; and the removal of exiting items. For entering items, I separately distinguish hospitals entering the product category, manufacturers entering the category, and new models introduced by existing manufacturers. The components sum to the change in the quantity-weighted price by construction.

Because hospitals enter and leave the ECRI reporting sample over time, changes in the composition of reporting facilities could affect this decomposition. I therefore repeat it using only hospitals observed throughout each event window. These hospitals account for approximately 68–90 percent of purchased quantities, depending on the cohort.

In this balanced-hospital sample, the average change in the price paid over years 0–3 remains negative (-0.22 log points, s.e. 0.12). The price change of continuing items is close to zero (-0.009, s.e. 0.051). Thus, the results continue to point toward changes in the products purchased rather than lower prices for the same products. These estimates use a different sample and comparison from the main patent-expiration event study, so their magnitudes are not directly comparable.

Competition and the Incidence of Production-Cost Savings

The distribution of production-cost savings between manufacturers and buyers may depend on competition. When manufacturers face limited competition, a reduction in production costs may increase their margins without substantially changing prices. In more competitive markets, lower production costs may allow manufacturers to reduce prices or win additional sales.

The firm-level results show that a 10-percentage-point increase in the production-cost patent score is associated with a 5.0-percentage-point increase in gross margin (Section 7). However, I cannot reliably estimate how this relationship varies with competition. Interactions between production-cost invention and market concentration are imprecise and change sign depending on whether I measure concentration using the Herfindahl index or the four-firm concentration ratio.

The ECRI data also do not support a precise comparison of transaction-price responses

across device markets with different levels of competition. Consequently, I cannot determine how much of the production-cost saving manufacturers retain or how this share varies with market structure. The evidence is consistent with manufacturers retaining some savings as margin, but the incidence of those savings remains uncertain.

Classifying Patents by Payment Setting

Section 7.4 examines how much device invention is associated with technologies that remain outside prospective payment. I measure this using the distribution of patents across payment settings, rather than the distribution of medical spending. This distinction matters because the price of a device is often included in a broader procedure or episode payment and cannot be separately observed in claims.

I use the patent-to-HCPCS crosswalk, containing approximately 3.4 million links, to assign device patents to HCPCS categories and classify those categories according to their current Medicare payment rules. The sample contains approximately 16,400 device patents per year. I weight patent links by their cosine similarity so that a patent associated with several billing codes contributes more to the codes with which its text is most closely related.

Under a narrow definition, I classify as outside prospective payment devices paid through the durable medical equipment, orthotics, and prosthetics fee schedules (HCPCS categories E, K, and L) and outpatient devices receiving separate pass-through reimbursement (relevant C codes). These categories account for 26.0 percent of the cosine-weighted patent flow.

A broader definition also includes supplies classified under HCPCS category A. Under this definition, the share increases to 44.0 percent. The difference reflects uncertainty about which supplies remain separately reimbursed rather than being included in larger payment units. I therefore treat the two figures as alternative classifications, rather than exact bounds on the share of invention outside prospective payment.

Neither definition fully captures devices used in physician offices. Some of these devices are paid through the practice-expense component of Medicare's physician fee schedule, rather than through a separate device payment. I use CPT place-of-service information to refine the classification where possible, but cannot assign every patented device to a unique payment setting.

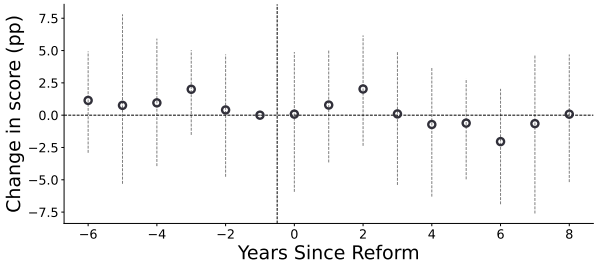
These estimates also do not measure the share of device spending exposed to prospective payment. Device costs included within hospital and physician payments are not separately observed, so the patent-to-billing-code crosswalk cannot establish what fraction of device expenditures remains under fee-for-service reimbursement. Industry spending estimates and MedPAC reports provide useful context, but do not resolve this measurement problem.

Finally, prospective payment already covers much of inpatient and outpatient care through the DRG system introduced in 1983 and OPPS introduced in 2000. The relevant policy comparison is therefore not a shift from no prospective payment to universal prospective payment. Instead, these classifications describe the share of device patenting associated with settings in which providers may still face different incentives to reduce device and care-delivery costs.

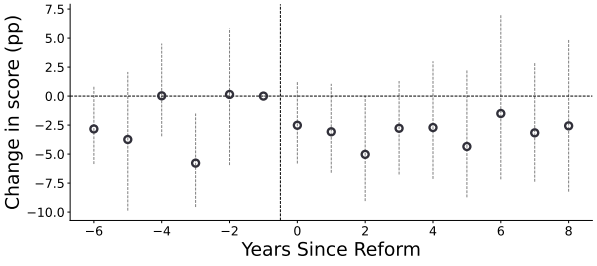
E Appendix Figures and Tables

Figure A.1: Event Studies for Reforms That Change the Returns to Innovation

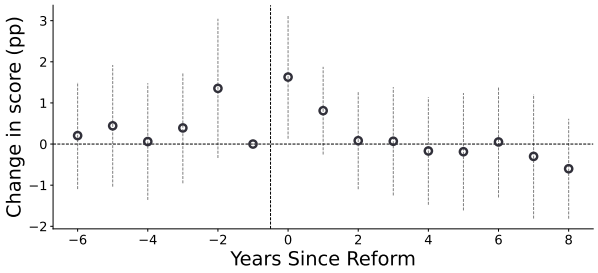
(a) FDA Class III to II: Production Cost



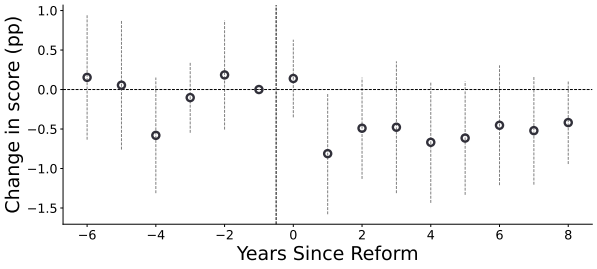
(b) FDA Class III to II: Episode Cost



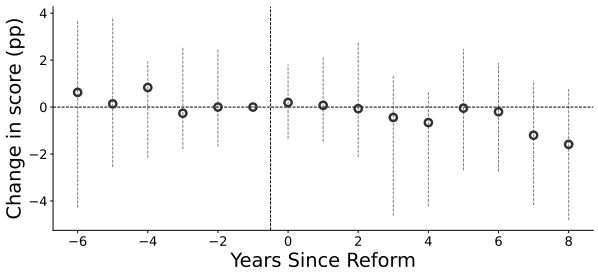
(c) FDA Class II to I: Production Cost



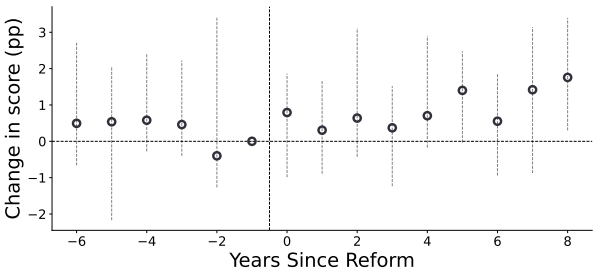
(d) FDA Class II to I: Episode Cost



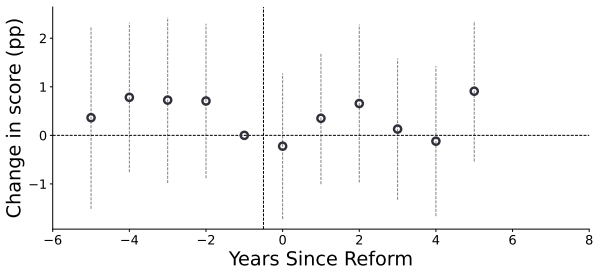
(e) DMEPOS Competitive Bidding: Production Cost



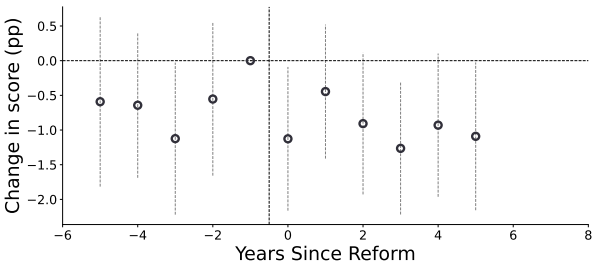
(f) DMEPOS Competitive Bidding: Episode Cost



(g) ACA Device Tax: Production Cost

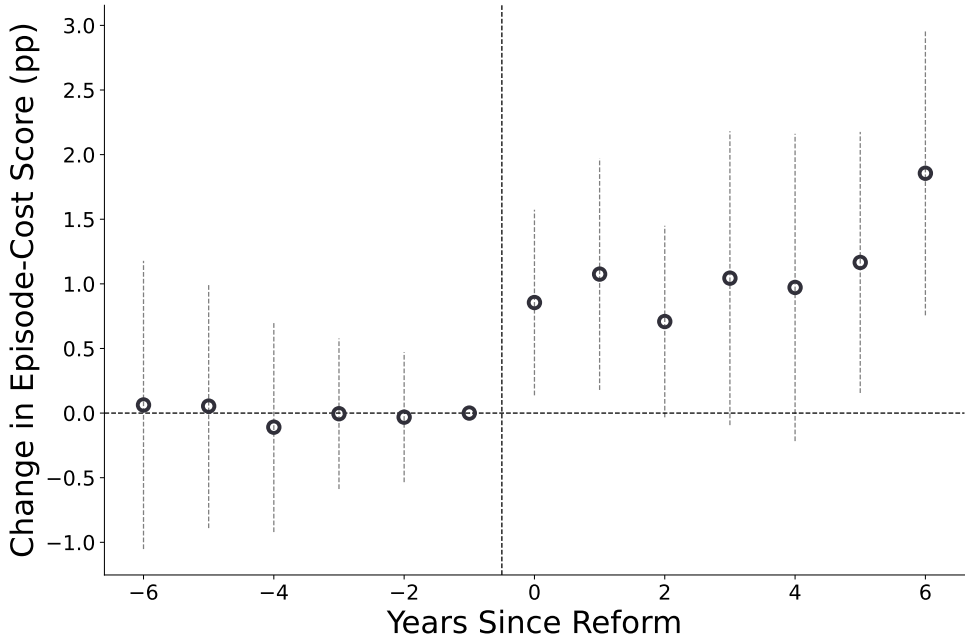


(h) ACA Device Tax: Episode Cost



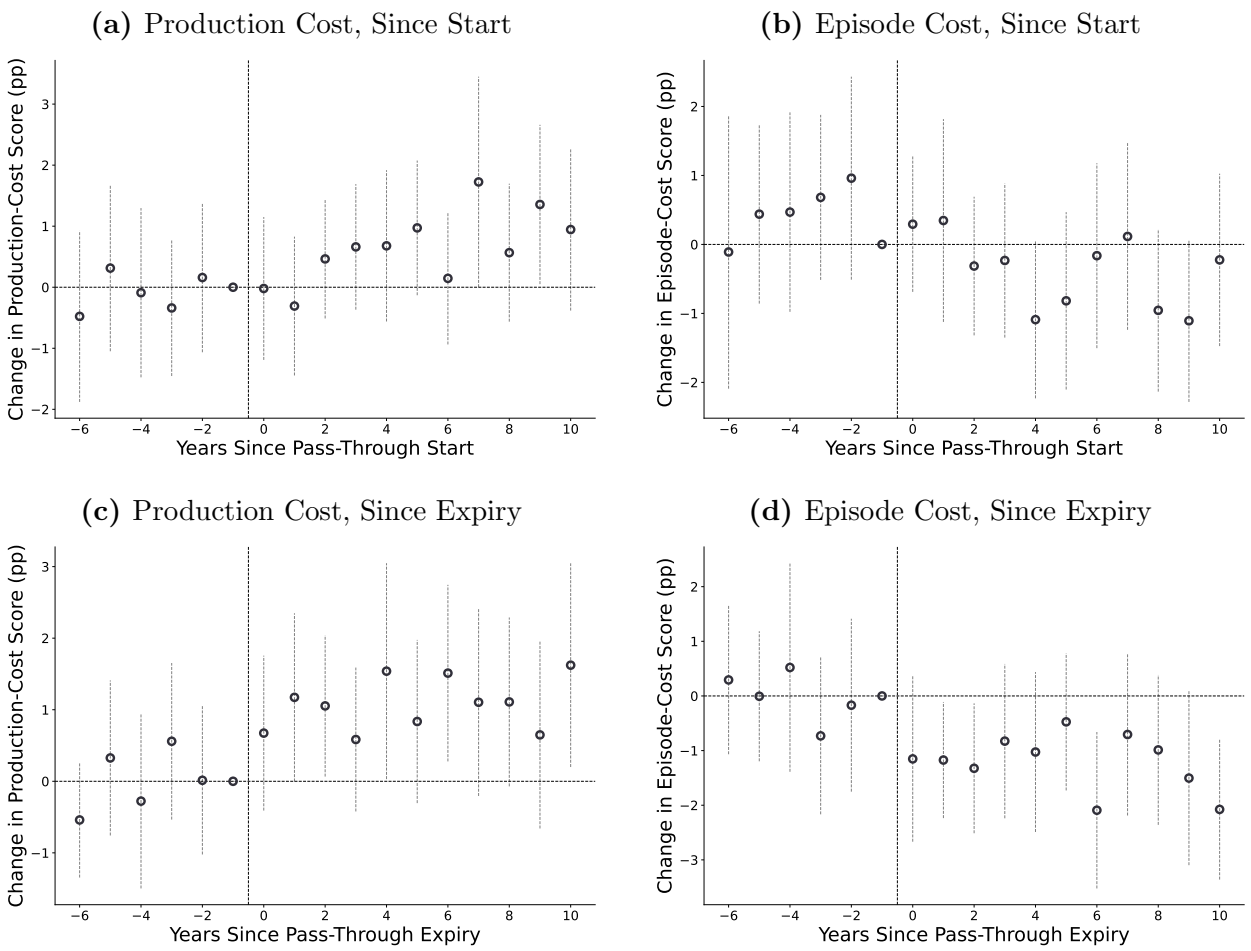
Note: Each panel plots event-study coefficients, relative to the year before the reform, with 95% confidence intervals, for the designs in Table 5. For FDA Class III to II (12 treated product codes) and DMEPOS (13 of 47 categories), intervals and joint pre-trend tests use a wild cluster bootstrap with the null imposed (Webb weights, 9,999 draws); intervals invert that test; for Class II to I and the device tax, standard errors are clustered by product code. Joint pre-trend p-values (production; episode): Class III to II 0.91; 0.07; Class II to I 0.51; 0.05; DMEPOS 0.95; 0.38; device tax 0.85; 0.39.

Figure A.2: Post-Acute Prospective Payment Event Study



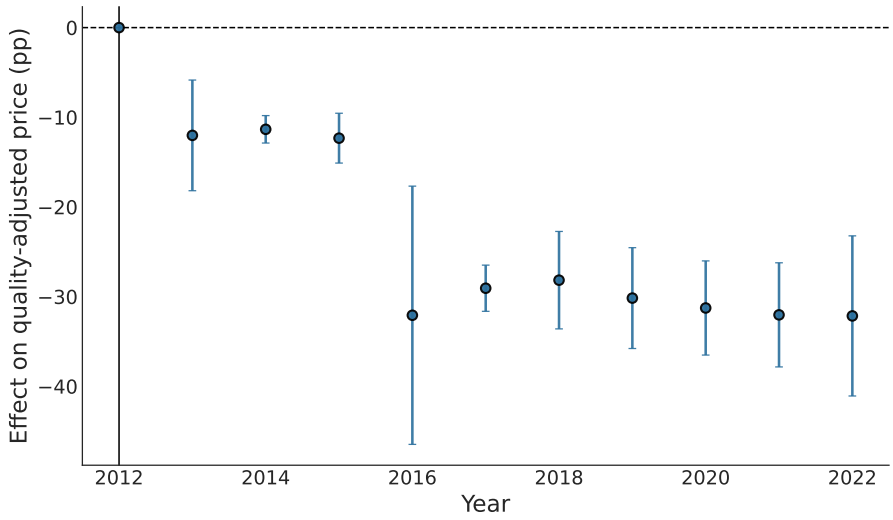
Note: Event-study coefficients for treated technology groups relative to matched comparison groups, by years since the reform (reference: the year before), with 95% confidence intervals clustered by CPC main group. The design pools the first payment reform in each post-acute setting that rewards an episode-cost channel: home health (2000), inpatient rehabilitation (2001), and long-term care hospitals (2002), each dated to the first rule publishing its payment design; 26 treated and 49 comparison groups, matched on pre-reform level and size, with dental groups excluded from the comparison pool. The outcome is the episode-cost score on the channels each payment rule rewards: complications, length of stay, and site of care for home health; complications and length of stay for rehabilitation and long-term care. Pooled estimate 1.11 percentage points (s.e. 0.48); the pre-reform coefficients are jointly zero ($p = 0.996$).

Figure A.3: Device Pass-Through: Event Studies by Years Since Start and Since Expiry



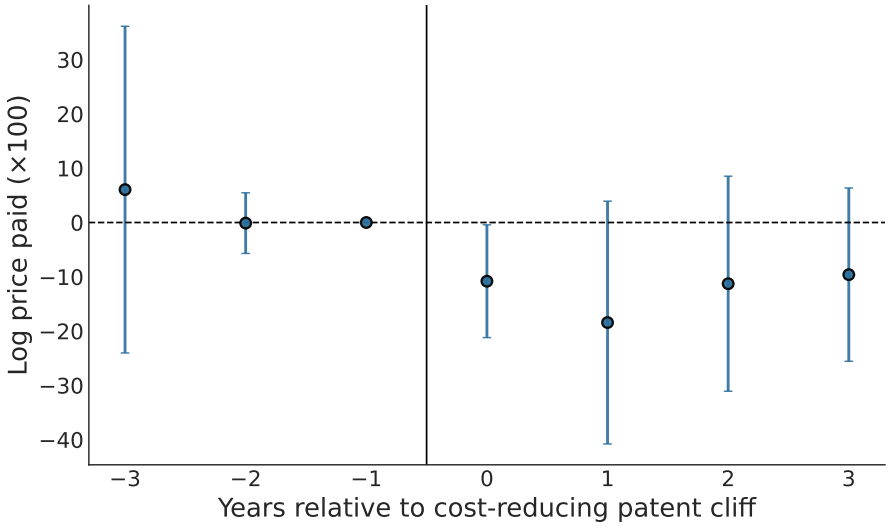
Note: The 86 device categories CMS granted pass-through payment between 2000 and 2018 (one code per patent), compared with codes that never received it; 80 categories are observed in more than one code-year cell and identify the estimates. *Panels (a)–(b)* plot coefficients by year relative to the start of pass-through payment, using the preceding year as the reference and stacking observations by start year. The pre-reform coefficients provide a test for differential trends underlying the expiry estimates in Table 4. They are jointly statistically insignificant ($p = 0.93$ for the production-cost score and $p = 0.78$ for the episode-cost score). Because pass-through payments last two to three years, later coefficients also capture the response to their expiration. The start of pass-through payment is not estimated separately in Tables 1 and 4, since the years receiving pass-through payment serve as the baseline for the expiry estimates. The production-cost score does not change significantly when pass-through begins (-0.05 percentage points, s.e. 0.30). *Panels (c)–(d)* plot coefficients by years since expiry, relative to the last shielded year, stacked by expiry year. Code and year fixed effects by cohort; 95% confidence intervals clustered by code.

Figure A.4: Quantity-Weighted Prices Fall Faster in Cost-Innovative Device Markets



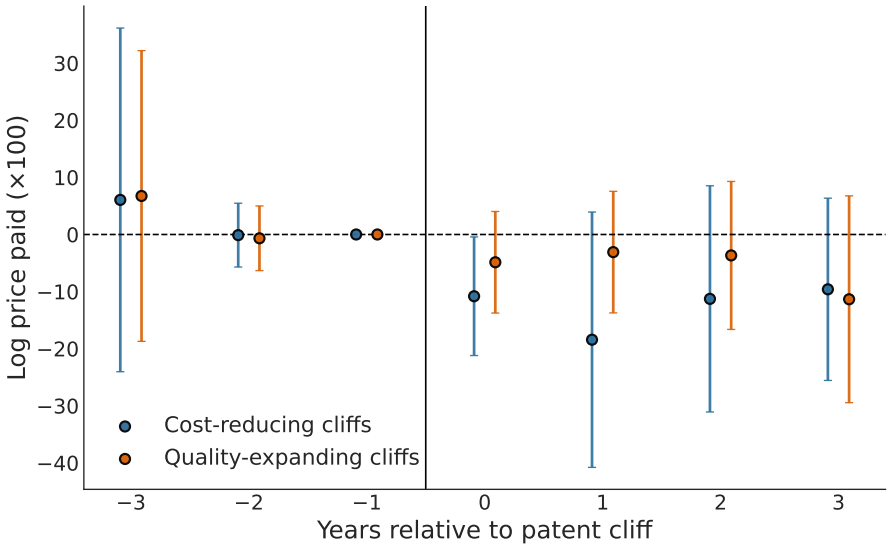
Note: Year-by-year interaction of a product code's cost-reducing innovation with the hedonic log-price index (ECRI 2012–2022). Innovation is per standard deviation, predetermined from patents filed by 2012, and linked through the patent-to-product-code embedding crosswalk. The index has 150 description-embedding controls and facility and product-code fixed effects; 95% confidence intervals are clustered by product code. The gap is statistically significant in every year and is robust to richer quality controls and to dropping the 2016 reporting anomaly. It comes from buyers shifting purchases toward cheaper-for-quality products; weighting products equally, it is null (Appendix D).

Figure A.5: Device Prices Around Cost-Reducing Patent Expirations



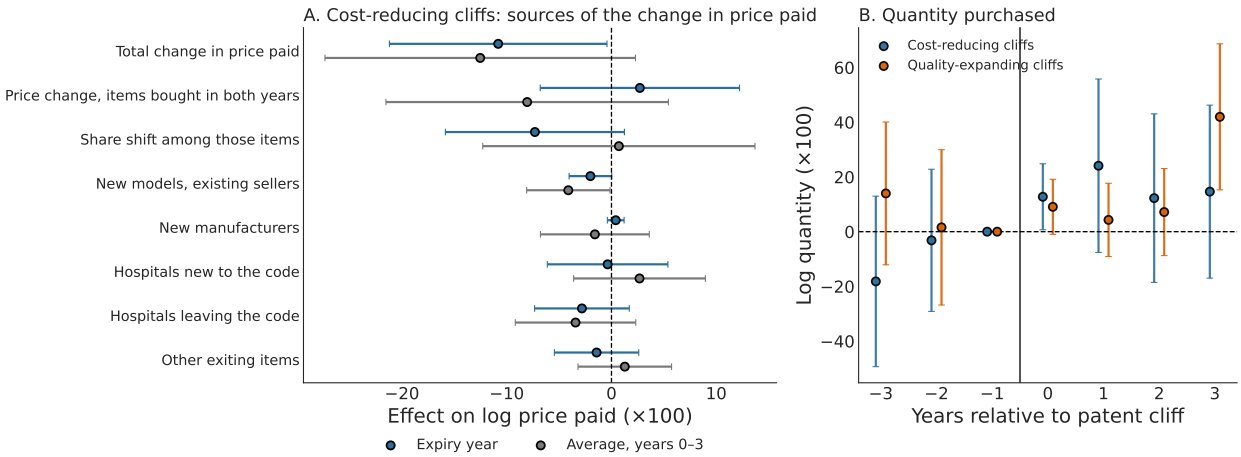
Note: Event study of the log quantity-weighted price hospitals pay in a product code around large expirations of cost-reducing patents (filing+20; 24 product codes). Each treated code is matched to three control codes (Mahalanobis distance on the prior year's price level, price change, quantity, and product count) drawn from codes with no large expiration of cost-reducing or quality-expanding patents within three years. Code-by-event and event-by-year fixed effects; reference year $t = -1$. Bars are 95% confidence intervals clustered by product code. The price is 0.074 log points lower in the expiration year (s.e. 0.059).

Figure A.6: Prices After Cost-Reducing and Quality-Expanding Patent Expirations



Note: Large expirations of cost-reducing patents (blue, 24 product codes) and of quality-expanding patents (orange, 21 codes), each matched to its own controls as in Figure A.5 and estimated jointly. In the expiration year the price falls 0.074 log points after cost-reducing expirations (s.e. 0.059) and 0.018 after quality-expanding expirations (s.e. 0.053). The difference is 0.055 (s.e. 0.062) in the expiration year and 0.084 (s.e. 0.079) averaged over years 0–3. Bars are 95% confidence intervals clustered by product code.

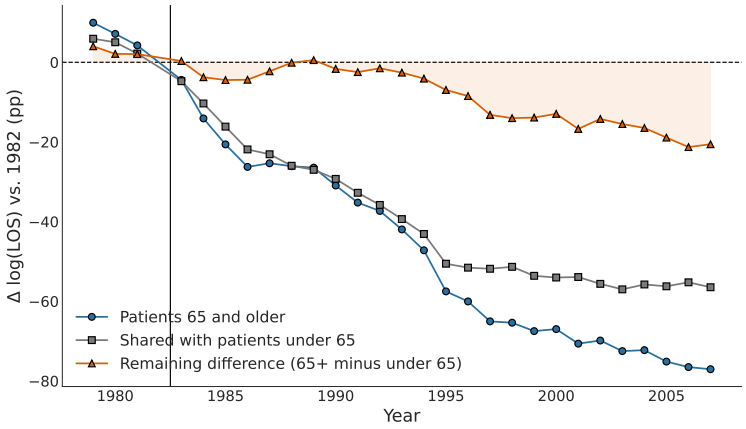
Figure A.7: Decomposing the Price Change After Patents Expire



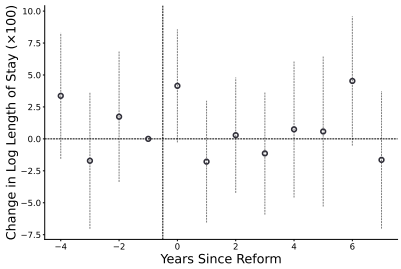
Note: Panel A splits the change in a product code's log quantity-weighted price after large cost-reducing expirations into the price change of items bought in both years (an item is a manufacturer's catalog number at a hospital), the shift in purchase shares among those items, entering items (from hospitals new to the code, manufacturers new to the code, and new models from existing sellers), and exiting items. The components sum exactly to the total. Blue: the expiration year; gray: the average of years 0–3. Panel B: log total quantity purchased in the code after cost-reducing (blue) and quality-expanding (orange) expirations. Design as in Figure A.5; 95% confidence intervals clustered by product code.

Figure A.8: Hospital Stays and Outcomes after the 1983 Reform (NHDS)

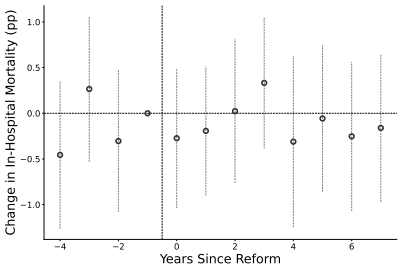
(a) Length of Stay by Age Group, Within Procedure



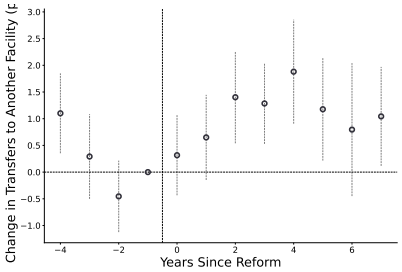
(b) Length of Stay: Medicare vs. Private



(c) In-Hospital Mortality: Medicare vs. Private

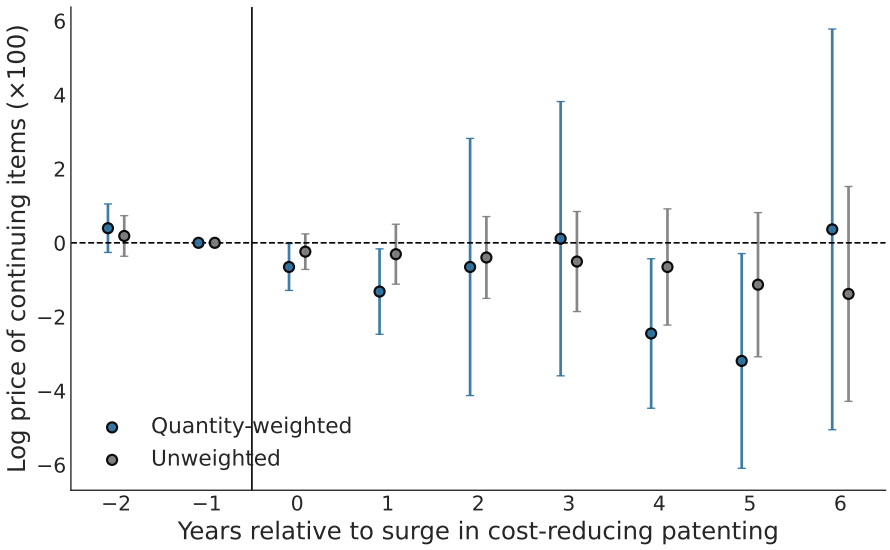


(d) Transfers: Medicare vs. Private



Note: *Panel (a)*: change in log length of stay relative to 1982, within procedure. The decline for patients 65 and older (blue) splits into the decline shared with patients under 65 (gray) and the remaining difference between the age groups (orange). The shared part is between about 72 and 100 percent of the decline from 1984 to 2007. Because the decline is the same for Medicare and privately insured patients of the same age (panel (b), and within five-year age bands; Section 7), the remaining difference reflects age rather than Medicare. *Panels (b)–(d)*: coefficients on Medicare × year, by years since the 1983 reform (reference: the year before), and 95% confidence intervals, from NHDS discharges of patients aged 55–74, 1979–1990. The comparison group is patients aged 55–74 whose expected payer is commercial insurance other than Blue Cross. Outcomes are computed by procedure, expected payer, and year. Regressions include procedure-by-payer and procedure-by-year fixed effects, are weighted by discharges, and are clustered by procedure. Pre-reform coefficients are jointly insignificant for stays ($p = 0.21$) and mortality ($p = 0.18$). The 1979 transfer coefficient is high, so the transfer pre-period is not jointly zero; excluding 1979, the pre-trend p is 0.13 and the estimate is +1.14 percentage points. NHDS changed its discharge-status codes in 1981; starting the sample in 1981, the estimate is +1.29 percentage points.

Figure A.9: Continuing Items Are Not Repriced After Surges in Cost-Reducing Patenting



Note: Stacked event study of the log price of continuing items (a manufacturer's catalog number at a hospital, relative to its own price in $t_0 - 1$) around years in which a product code's cost-reducing patent flow (linked patents times their mean cost score) is at least one standard deviation above the code's own mean (390 codes; surge years 2013–2019). Controls are items in codes with no surge in cost-reducing or quality-expanding patenting from $t_0 - 3$ to $t_0 + 6$ (137 codes). Item-by-event and event-by-year fixed effects; prices 2012–2022; item-events with any annual change above 0.5 log points dropped. Blue: weighted by the item's quantity in $t_0 - 1$; gray: unweighted. Bars are 95% confidence intervals clustered by product code.

Table A.1: Reforms Considered and Not Used

Reform (year)	What it changed	Why not used
Start of device pass-through (2000–18)*	Medicare paid hospitals separately for a new device's price for two to three years	Not a separate test: it is the baseline of the pass-through expiry design; production-cost direction does not change when it begins (−0.05 percentage points, s.e. 0.30)
Outpatient prospective payment (BBA 1997; launched 2000)*	Packaged established devices into fixed procedure payments	Devices first paid after 1996 whose cost was not insignificant were paid separately for two to three years, so packaging reached new devices only when that separate payment expired, which the expiry design tests. Adoption leaves three years before the first expirations, too short to separate a response to packaging from the pass-through cycle; estimated directly, it does not change production-cost direction (−0.3 percentage points, s.e. 0.4, flat pre-trend)
Outpatient packaging expansions (2008; comprehensive APCs 2015)*	2008: packaged supportive services (image guidance, imaging supervision and interpretation, intraoperative services) into the procedures they accompany; 2015: one payment for a device-intensive procedure and everything billed with it	2008: too few affected device groups for inference; 2015: affects the joint-replacement procedures that BPCI and CJR also cover, with few post-reform years
Inpatient new-technology add-on payments (BIPA 2000; first approved technology FY2003)*	Paid hospitals part of a new technology's cost above the DRG for two to three years	Few technologies a year, each with its own dates; an early test at the CPC-subclass level was uninformative (7 subclasses, 3 treated); not rebuilt at the technology level
Skilled nursing facility prospective payment (1998)*	Per-diem payment	A per-diem payment does not reward shorter stays or fewer complications; not treated as exposed in the post-acute design
Comprehensive Care for Joint Replacement (2016)*	Mandatory joint-replacement bundles in selected areas	Episodes already bundled by BPCI; its effect is partly contained in the BPCI estimate
Hospital-Acquired Condition (2008) and Readmissions Reduction (2012) penalties*	Payment reductions for poor outcomes	Do not make providers bear a device's cost; the technologies they target are those the post-acute reforms treated a few years earlier, so their pre-reform periods are not clean
Medicare Shared Savings Program (accountable care organizations, 2012)*	Shared savings on total spending, voluntary	Technology exposure can only be measured through Medicare spending, which coincides with exposure to the 1983 reform and BPCI; cost-reducing invention in exposed fields was already rising before the program
Capital prospective payment (1992)*	Folded capital costs into DRG payments	Coincides with the 1990–91 squeeze; discussed in Section 6.2 as a competing explanation
Introduction of Medicare (1965)*	Extended cost-based reimbursement to the elderly	Cost-based reimbursement by private insurers predated Medicare, so it did not change which costs providers bore; no break in direction
ESRD prospective payment (2011)	Folded dialysis drugs into an already-prospective bundle	Dialysis devices were bundled under the 1983 composite rate, so the reform has no device-side first stage
Medicare Advantage (capitation)	Paid private plans a fixed amount per enrollee	No direct map from patents to the technologies it affects; enrollment changed gradually

Note: Payment and regulatory reforms that could plausibly change the direction of device invention but are not used in the main analysis. *Estimated at some stage of this project; the reason summarizes the result. Sources: Balanced Budget Act of 1997, sec. 4523; Balanced Budget Refinement Act of 1999, sec. 201 (42 U.S.C. 1395l(t)(6)); CY2008 OPPS final rule (72 FR 66580); CY2015 OPPS final rule (79 FR 66770); Benefits Improvement and Protection Act of 2000, sec. 533; Deficit Reduction Act of 2005, sec. 5001(c); Affordable Care Act, secs. 3022 and 3025.

Table A.2: Technology Groups in the 1983 Design

CPC subclass	Inpatient hospital-technology fields	Comparison device fields
A61B: Diagnosis; surgery	A61B17 surgical instruments; A61B18 energy-based surgery; A61B34 robotic and computer-aided surgery; A61B42 surgical gloves; A61B46 surgical drapes; A61B50 holders for surgical instruments; A61B90 other surgical accessories	A61B3 eye examination; A61B9 percussion instruments; A61B13 tongue depressors
A61F: Implants, prostheses, orthoses, dressings	A61F2 prostheses and implants	A61F3 leg-lengthening pieces; A61F4 aids for disabled persons to operate devices; A61F5 orthopaedic and nursing devices; A61F6 contraceptive devices; A61F17 first-aid kits
A61G: Patient transport and accommodation	A61G1 stretchers; A61G7 nursing beds and patient lifts; A61G9 bedpans and urinals; A61G10 treatment rooms; A61G11 infant incubators; A61G12 hospital nursing equipment; A61G13 operating tables	A61G3 patient transport vehicles; A61G5 wheelchairs
A61H: Physical therapy; massage	–	A61H1 passive exercise devices; A61H3 walking aids; A61H5 eye exercisers; A61H11 massage belts; A61H13 gum massage; A61H21 body-cavity massage; A61H35 baths for parts of the body; A61H36 sweating suits; A61H37 massage accessories; A61H39 reflex-point therapy
A61J: Medical containers; feeding devices	A61J15 feeding tubes	A61J7 oral medicine dispensers; A61J13 nipple shields
A61L: Sterilisation; materials for medical devices	A61L17 suture materials; A61L24 surgical adhesives; A61L27 prosthesis materials; A61L29 catheter materials; A61L31 other surgical-article materials; A61L33 antithrombogenic coatings	A61L12 contact-lens disinfection; A61L28 colostomy materials
A61M: Devices introducing media into the body	A61M16 ventilators and tracheal tubes; A61M25 catheters; A61M27 wound drainage; A61M39 medical tubes, connectors, and valves; A61M60 blood pumps	A61M15 inhalers; A61M21 sleep and consciousness devices; A61M35 devices applying remedies to the body
A61N: Electrotherapy; radiation therapy	–	A61N2 magnetotherapy; A61N5 radiation therapy; A61N7 ultrasound therapy

Note: The 1983 design assigns exposure by a patent's CPC main group. It compares 27 inpatient hospital-technology fields (24,200 patents filed 1977–1995) with 30 device fields used mainly outside inpatient care (8,129 patents), classified from their official titles before estimation. Drug fields, dental comparison fields, and 48 fields whose use spans inpatient and outpatient care are excluded. Each main group is followed by a short label written from its official CPC title; the full titles, each group's patents, and its continuous inpatient share (the robustness exposure) are in the replication files. The other designs link patents to reforms through billing codes or product codes rather than patent classes (Table A.5).

Table A.3: Technology Groups in the Post-Acute Design

CPC subclass	Treated groups [reform]	Matched comparison groups
A61B: Diagnosis; surgery	A61B5 diagnostic measurement [LT]	A61B1 endoscopes; A61B3 eye examination; A61B7 stethoscopes; A61B8 ultrasound diagnosis; A61B9 percussion instruments; A61B10 body-sample instruments; A61B13 tongue depressors; A61B16 autopsy devices; A61B18 energy-based surgery; A61B34 robotic and computer-aided surgery; A61B42 surgical gloves; A61B46 surgical drapes; A61B50 holders for surgical instruments; A61B90 other surgical accessories
A61F: Implants, prostheses, orthoses, dressings	A61F5 orthopaedic and nursing devices [IR]; A61F13 bandages and dressings [HH]	A61F9 eye treatment devices; A61F11 ear treatment devices; A61F17 first-aid kits
A61G: Patient transport and accommodation	A61G1 stretchers [LT]; A61G5 wheelchairs [IR]; A61G7 nursing beds and patient lifts [LT]	A61G3 patient transport vehicles; A61G10 treatment rooms; A61G11 infant incubators; A61G12 hospital nursing equipment; A61G13 operating tables
A61H: Physical therapy; massage	A61H1 passive exercise devices [IR]; A61H3 walking aids [IR]; A61H5 eye exercisers [IR]; A61H9 pneumatic massage [IR]; A61H11 massage belts [IR]; A61H13 gum massage [IR]; A61H21 body-cavity massage [IR]; A61H31 chest compression [IR]; A61H35 baths for parts of the body [IR]; A61H36 sweating suits [IR]; A61H37 massage accessories [IR]; A61H39 reflex-point therapy [IR]; A61H99 other physical therapy [IR]	–
A61J: Medical containers; feeding devices	A61J15 feeding tubes [LT]	A61J1 medical containers; A61J3 devices for forming medicine doses; A61J13 nipple shields; A61J19 spittoons
A61L: Sterilisation; materials for medical devices	A61L2 sterilisation [LT]; A61L15 dressing materials [HH]; A61L26 bandage materials [HH]	A61L9 air disinfection; A61L11 medical waste treatment; A61L12 contact-lens disinfection; A61L17 suture materials; A61L24 surgical adhesives; A61L27 prosthesis materials; A61L28 colostomy materials; A61L29 catheter materials; A61L31 other surgical-article materials; A61L33 antithrombogenic coatings
A61M: Devices introducing media into the body	A61M15 inhalers [LT]; A61M16 ventilators and tracheal tubes [LT]; A61M25 catheters [HH]	A61M3 syringes and irrigators; A61M13 insufflators; A61M19 local anaesthesia; hypothermia; A61M21 sleep and consciousness devices; A61M29 dilators; A61M31 devices placing remedies in body cavities; A61M35 devices applying remedies to the body; A61M37 other drug-delivery devices; A61M60 blood pumps
A61N: Electrotherapy; radiation therapy	–	A61N1 electrotherapy; A61N2 magnetotherapy; A61N5 radiation therapy; A61N7 ultrasound therapy

Note: The post-acute design assigns exposure by a patent's CPC main group. Each treated group enters in the first post-acute reform whose payment rewards it: [HH] home health (2000), [IR] inpatient rehabilitation (2001), [LT] long-term care hospitals (2002). The 49 comparison groups are matched to the 26 treated groups on pre-reform level and size. Each main group is followed by a short label written from its official CPC title; full titles are in the replication files.

Table A.4: Classifier Performance on Held-Out Patents

Channel	Positives	Base rate	AUC	Correlation
Production-cost-reducing	128	0.04	0.93	0.50
Quality-expanding	267	0.50	0.79	0.47
Episode-cost-reducing (composite)	54	0.14	0.92	0.70
Complication-reducing	177	0.37	0.81	0.59
Length-of-stay-reducing	48	0.14	0.95	0.75
Site-shifting	60	0.12	0.84	0.54
Labor-saving	145	0.22	0.77	0.53

Note: The deployed TF-IDF logistic classifiers (title, abstract, and summary of invention), evaluated on the 624 validation patents outside the training sample (643 drawn, 19 dropped for overlap with the training set), against labels from the deployed labeling model. A patent is a positive when its label is at least 0.5; the episode composite averages the complication, length-of-stay, and site-shifting labels. Positives counts the held-out patents with a positive label (unweighted); base rates, AUCs, and correlations use sampling weights that make the validation sample representative of device patents. Correlation is between the classifier score and the graded label.

Table A.5: Specification and Inference by Design

Estimated designs only; ESRD and Medicare Advantage are listed in Table A.1 and not estimated. Patents are dated by filing year. The outcome is the patent-text score (episode-cost or production-cost) that each design uses (Section 4), in the treated unit-year; level results use the log patent count. "Treated" and "post" refer to the exposed unit and the treatment date in the Reform column.

Reform	**Obs. level (N)**	**Patent → unit link**	**Estimator**	**FE + controls**	**Cluster (count)**	**Inference**
Inpatient PPS (1983)	patent (57 CPC groups; 32,328 pat, 1977–95)	CPC main group classified from its official title (no crosswalk)	binary DiD, two post windows (1983–90, 1991–95)	group + filing year	CPC group (56)	cluster-t
Post-acute PPS	patent × reform stack (131,525; 26 treated, 49 control main groups)	CPC main group → supplies the setting's payment covers	matched stacked DiD, pooled over three reforms	stack×group + stack×year	CPC main group	cluster-t
BPCI (2011)	patent (74,780; 2005–2020)	any of the top five procedure links (cosine ≥0.45) on a BPCI-covered code	binary-exposure DiD, t_0=2011	technology class + filing year	technology class	cluster-t
OPPS pass-through expiry	code × year × stack cells (446,154; 80 pass-through categories vs. 6,888 never-pass-through codes)	one top-cosine patent→HCPCS link	stacked DiD by start year; exposed minus shielded years, CMS dates	code×stack + year×stack	HCPCS code	cluster-t
FDA reclass III→II	product code × year (1,118; 12 FDA-initiated events, 22 controls)	FDA replication patent→product-code link	matched stacked DiD, each event with its own two controls	PC×stack + stack×year	product code (34)	wild-cluster bootstrap
FDA reclass II→I	product code × year (10,835; 129 treated, 136 controls)	FDA replication patent→product-code link	matched stacked DiD	PC×stack + stack×year	product code (265)	cluster-t
DMEPOS bidding	category × year (8,613; 13 treated / 34 control categories, 1996–2016)	DME paper's patent→category link	stacked DiD (2004/07/11/12)	category×stack + stack×year	DME category (47)	wild-cluster bootstrap

Table A.5 (continued)

Reform	**Obs. level (N)**	**Patent → unit link**	**Estimator**	**FE + controls**	**Cluster (count)**	**Inference**
ACA excise tax	PC × year (6,301; 515 tr/159 ctrl, matched; '05–'15)	cosine patent→product-code; IRS-taxable vs. exempt	Mahalanobis-matched DiD	PC + year	product code (674)	cluster-t
Prescription-to-OTC	molecule × year (1,744; 24 switches)	Orange Book NDA, cosine ≥0.45	stacked DiD, K=5 Mahalanobis, ±6 yr	stacked	switch (24)	cluster-t; timing checks

Table A.6: Each Reform's Effect on Every Cost Channel

Reform	Production	Individual episode-cost channels				Episode-cost aggregations			Quality
		Complications	Length of stay	Site of care	Labor†	Clinical	Mean-3	Mean-4	
A. Reforms That Change Who Bears the Costs									
BPCI bundled payment (2011)	−0.41	0.73	2.33***	1.69***	2.23***	1.53**	**1.58***	1.74***	−1.31*
Inpatient PPS (1983), 1991–95	−2.46**	2.74	3.87**	−0.47	3.03*	**3.31***	2.05+	2.30*	−1.66
Post-acute PPS‡	−0.44	1.55**	0.61	1.34*	−0.42	1.08*	1.16*	0.77*	−0.67
Pass-through expiry	**1.02****	−0.66	−1.20*	−0.89*	−0.48	−0.93*	**−0.91***	−0.80*	0.07
B. Within Firm and Year									
BPCI bundled payment (2011)	−0.53	1.18*	2.63***	1.10**	0.84	1.90***	**1.64***	1.44***	−0.90
Pass-through expiry	**0.63***	−0.29	−0.89*	−0.44	−0.68		−0.54		0.01
Firm portfolio	0.17	0.24	0.03	−0.16	0.40+		0.03		−0.60+
C. Reforms That Change the Returns to Innovation									
FDA reclassification, Class III to II	**−0.96**	−2.26	1.23	−0.70	−0.61		**−0.57**		2.02
FDA reclassification, Class II to I	**−0.76**	−0.49	−0.32	−0.71	−1.14		**−0.51**		1.44+
DMEPOS competitive bidding	**−0.83**	0.31	0.50	0.59	−0.60		**0.47**		0.25
ACA device tax	**−0.22**	−0.37	−0.49+	−0.31	−0.76+		**−0.39**		0.29

Note: Each row re-estimates a design from Tables 4 and 5 with the outcome in each column; the design, units, fixed effects, clustering, and exposure scaling are the row's own, and only the outcome changes. Each cell is the effect in percentage points; **bold** cells are the estimates those tables report. The production-cost score is the probability that an invention lowers the device's manufacturing cost; the complication, length-of-stay, and site-of-care scores are the probabilities that it lowers the cost of an episode of care through that channel; Clinical is the mean of complications and length of stay, Mean-3 adds site of care, and Mean-4 adds labor saving; the quality score is the probability that it expands what the device does. The 1983 row is the 1991–95 coefficient of the baseline design. Clinical and Mean-4 are reported for the designs whose episode outcome they bracket; for the other rows, Mean-3 is the design's episode-cost outcome. Where a design matches comparison units on its primary outcome (the device tax and post-acute payment), the matched sample is held fixed across columns. Outside the bold cells, the significant effects of reforms that change who bears the costs are concentrated in the individual channels of the episode-cost score, the production-cost score, which moves in the opposite direction (Figure 3), and labor saving, which moves in the same direction as the episode-cost score. None of the estimates for reforms that change the returns to innovation are significant at the 5 percent level. Standard errors are clustered on the design's treatment unit, except the FDA Class III-to-II and DMEPOS rows, which use a wild cluster bootstrap (12 treated product codes; 13 of 47 categories). †Labor saving has no counterpart in realized outcomes, so it is not part of any primary outcome. ‡The post-acute baseline follows each setting's payment rule (Mean-3 for home health; Clinical for rehabilitation and long-term care), which no single column matches; it is 1.11 (s.e. 0.48). +, *, **, and *** denote statistical significance at the 0.10, 0.05, 0.01, and 0.001 levels, respectively.

Table A.7: Alternative Comparison Groups and Technology-by-Year Effects, by Design

	Baseline (1)	Alternative (2)	Alternative comparison group	Subclass × Year (3)
A. Reforms That Change Who Bears the Costs				
Inpatient PPS (1983), 1991–95	3.31*	2.47*	all other fields, 1976–95	4.00***
	(1.29)	(1.17)		(0.74)
Inpatient PPS (1983), 1983–90	1.03	0.95	all other fields, 1976–95	2.02**
	(0.72)	(0.65)		(0.62)
BPCI bundled payment (2011)	1.58***	1.29**	every procedure-linked patent	1.67**
	(0.46)	(0.44)		(0.59)
Post-acute PPS	1.11*	0.98+	broader matched pool	1.10*
	(0.48)	(0.50)		(0.42)
Pass-through expiry, production	1.02**	1.02**	codes not yet billed in 2000	0.85**
	(0.35)	(0.35)		(0.32)
Pass-through expiry, episode	−0.91*	−0.92*	codes not yet billed in 2000	−0.82*
	(0.37)	(0.39)		(0.37)
B. Reforms That Change the Returns to Innovation				
FDA Class III to II, production	−0.96	−0.89	matched controls pooled across events	–
	(1.87)	(1.83)		
FDA Class III to II, episode	−0.57	−0.51	matched controls pooled across events	–
	(2.32)	(1.73)		
FDA Class II to I, production	−0.76	−0.76	matched controls pooled across events	–
	(0.58)	(0.69)		
FDA Class II to I, episode	−0.51	−0.46	matched controls pooled across events	–
	(0.45)	(0.42)		
DMEPOS competitive bidding, production	−0.83	−0.38	alternative patent-to-category link	–
	(0.70)	(0.62)		
DMEPOS competitive bidding, episode	0.47	−0.14	alternative patent-to-category link	–
	(0.36)	(0.36)		
Device tax, production	−0.22	0.61*	unmatched full pool	–
	(0.35)	(0.29)		
Device tax, episode	−0.39	–	not estimated	–
	(0.24)			

Note: Each cell is the reform's effect on a patent-text score in percentage points, with standard errors in parentheses. Column (1) is the baseline estimate from Table 4 or Table 5, whose event studies carry each design's pre-trend test. Column (2) uses the alternative comparison group named in the next column. Column (3) adds technology-subclass-by-year fixed effects to the baseline, so each technology subclass has its own year effects; it is reported for the reforms that change who bears the costs. A BPCI patent is exposed if any of its procedure links falls on a code covered by a BPCI episode. Significance for the FDA Class III-to-II and DMEPOS baselines uses a wild cluster bootstrap. +, *, **, and *** denote statistical significance at the 0.10, 0.05, 0.01, and 0.001 levels, respectively.

Table A.8: The Payment Unit Providers Bear under Each Reform

Reform	Payment unit provider bears	Approx. dollar magnitude
Inpatient PPS (1983)	inpatient admission	several $1,000s per discharge
OPPS pass-through expiry (2000–18)	single outpatient procedure (APC)	$100s to a few $1,000s
BPCI (2011)	90-day joint-replacement episode	$27,265–30,551 per episode
Post-acute PPS	per-stay / episode	$16,649 per case (IRF); $3,000–40,000 range

Note: The payment unit whose cost each reform places on the provider, with a representative Medicare payment. The BPCI figure is a program-evaluation mean (Dummit et al., 2016), and the inpatient-rehabilitation figure is from MedPAC. The DRG and OPPS amounts are representative figures from the CMS FY1984 inpatient and CY2000 outpatient prospective-payment rules. Dollars are nominal and from different years. The size of the unit spans more than an order of magnitude regardless of inflation: a single outpatient procedure $<$ an inpatient admission $<$ a post-acute stay $<$ a 90-day episode. The effect on direction does not scale with these units (Section 6.3).

Table A.9: Within-Technology and Within-Firm Composition

	1983 PPS	BPCI	Post-acute	Pass-through expiry	
Specification	episode	episode	episode	production	episode
Baseline (Table 4)	3.31*	1.58***	1.11*	1.02**	−0.91*
	(1.29)	(0.46)	(0.48)	(0.35)	(0.37)
CPC subgroup FE (finest code)	1.34+	1.31***	0.55	0.74*	−0.53+
	(0.68)	(0.28)	(0.38)	(0.30)	(0.32)
Pre-reform subgroup mix (reweighted)	1.51+	1.33***	0.61+		
	(0.84)	(0.34)	(0.36)		
Company-assigned patents	4.29*	1.60***	1.18*		
	(1.71)	(0.48)	(0.48)		
Incumbent firms	2.68+	1.74***	0.87		
	(1.51)	(0.47)	(0.57)		
Excluding post-reform entrants	3.42**	1.65***	0.85+		
	(1.19)	(0.47)	(0.49)		

Note: Estimates in percentage points, standard errors in parentheses. CPC subgroup FE adds fixed effects for each patent's finest CPC code. Pre-reform subgroup mix reweights each class's subgroups to their pre-reform shares (not computed for pass-through, where 58 percent of post-reform patents fall in codes with no pre-reform patents). Company-assigned patents drops patents assigned to individuals; incumbent firms keeps firms that patented before the reform; excluding post-reform entrants drops firms whose first patent is after it. +, *, **, and *** denote statistical significance at the 0.10, 0.05, 0.01, and 0.001 levels, respectively.

Table A.10: Effects in Standard Deviations of Normal Change

Design	Coef. (×100)	Units / changes	SD of placebo changes	Effect / SD		
				Primary [95% CI]	Raw	Unwtd.
Panel A. Reforms that change who bears a cost						
Inpatient PPS (1983), 1991–95	3.31 (1.29)	22 / 275	2.18	1.52 [0.36, 2.68]	1.39	1.43
production cost	−2.46 (0.85)	22 / 275	2.57	−0.96 [-1.61, −0.31]	−0.80	−0.92
BPCI (2011)	1.58 (0.46)	95 / 980	1.77	0.89 [0.38, 1.40]	0.74	0.88
production cost	−0.41 (0.26)	95 / 980	2.22	−0.19 [-0.42, 0.04]	−0.14	−0.19
Post-acute PPS (pooled)	1.11 (0.48)	29 / 197	1.15	0.97 [0.15, 1.78]	0.82	0.59
production cost	−0.44 (0.46)	29 / 197	1.15	−0.39 [-1.18, 0.40]	−0.32	−0.25
Pass-through expiry, production	1.02 (0.35)	253 / 1940	1.26	0.81 [0.27, 1.35]	0.52	0.76
episode	−0.91 (0.37)	253 / 1940	1.51	−0.60 [-1.09, −0.12]	−0.44	−0.57
Panel B. Reforms that change only the size of the payoff						
FDA Class III to II	−0.96 (1.87)	16 / 227	1.84	−0.52 [-2.51, 1.46]	−0.38	−0.39
FDA Class II to I	−0.76 (0.58)	94 / 1627	2.11	−0.36 [-0.90, 0.18]	−0.33	−0.27
DMEPOS competitive bidding	−0.83 (0.70)	24 / 47	2.12	−0.39 [-1.05, 0.26]	−0.38	−0.25
Medical device tax	−0.22 (0.35)	52 / 601	1.31	−0.17 [-0.68, 0.35]	−0.09	−0.17

Note: Coef.: the main estimate in percentage points of the score, standard error in parentheses (1983, BPCI, and post-acute: treated indicator; pass-through: exposed minus shielded years). Comparison changes: for every unaffected unit and every start year whose windows fit in 1976–2009 (pass-through: 1990–2020), the change in the unit's mean score between the design's pre- and post-reform windows, net of unit and filing-year means, with at least 30 patents per window. The 1983 windows are 1977–82 and 1991–95, eight years apart, and the units are device fields classified as used mainly outside inpatient care, including the dental fields the estimate leaves out of its comparison group. For the post-acute reforms, the outcome follows each setting's payment rule, so normal change is measured on the same mix of channels in the matched comparison groups. Net: the variance of the changes minus the average sampling variance of the two window means. Weighted: each change weighted by the smaller window's patent count. SD of placebo changes: the net, weighted standard deviation used for the primary ratio. Windows (years before/after): 1983, 6/5 with an eight-year gap; BPCI, 6/10; post-acute, 6/7; pass-through, 3/7; FDA and DMEPOS, 10/10; device tax, 5/6. Raw: the ratio using the raw rather than net standard deviation; Unwtd.: using unweighted changes. The 95 percent interval holds the benchmark fixed. Panel B: units are the design's never-treated product codes or categories; FDA and DMEPOS compare all years before a reform with all years after it, so their windows are capped at ten years on each side; the device tax uses its own five pre-reform and six post-reform years. The FDA Class III-to-II estimate is imprecise.

Table A.11: Post-Acute Estimates with One Design Rule Removed at a Time

	Estimate	s.e.	Pre-trend p	Treated groups	Comparison groups
Adopted design	1.11*	(0.48)	1.00	26	49
Home-health links not limited to items the episode payment covers	1.03*	(0.43)	0.66	27	49
Skilled nursing (per-diem payment) treated as exposed	0.50	(0.52)	0.05	26	49
Dressing materials and disinfection (A61L) not assigned	1.16*	(0.51)	0.99	23	46
Home health dropped	1.36**	(0.47)	0.87	22	49

Note: Each row re-estimates the post-acute design of Table 4 (payment-rule outcome, percentage points; standard errors clustered by CPC main group) with one of its rules removed and everything else, including the matching of comparison groups, unchanged. Pre-trend p tests that the event-study coefficients before the reform are jointly zero. Absorbent articles (A61F13/15 and higher) are excluded from every design by the consumer-product rule, so the post-acute exclusion of them has no separate effect. +, *, **, and *** denote statistical significance at the 0.10, 0.05, 0.01, and 0.001 levels, respectively.

Table A.12: Estimates under a Stricter Consumer-Product Exclusion

	Baseline exclusion		Stricter exclusion	
	Estimate	Pre-trend p	Estimate	Pre-trend p
Inpatient PPS (1983), 1983–90	1.03 (0.72)	0.15	1.05 (0.72)	0.04
Inpatient PPS (1983), 1991–95	3.31* (1.29)	0.15	3.35* (1.30)	0.04
Bundled payment (BPCI, 2011)	1.58*** (0.46)	0.97	1.59*** (0.46)	0.97
Post-acute PPS	1.11* (0.48)	1.00	1.29* (0.51)	0.96
Pass-through expiry, production cost	1.02** (0.35)	0.93	1.02** (0.34)	0.93
Pass-through expiry, episode cost	−0.91* (0.37)	0.78	−0.91* (0.38)	0.77
FDA reclassification III→II	−0.96 (1.87)		−0.96 (1.86)	
FDA reclassification II→I	−0.76 (0.58)		−0.76 (0.58)	
DMEPOS competitive bidding	−0.83 (0.70)		−0.81 (0.70)	
Medical device tax	−0.22 (0.35)	0.85	−0.26 (0.35)	0.85

Note: Each row is a baseline estimate from Table 4 or Table 5 (percentage points, standard error in parentheses) under the consumer- and non-human-health-product exclusion used throughout (columns 1–2) and under a stricter list (columns 3–4) that adds 14 CPC groups (additional sexual-aid groups; anti-snoring and nasal dilators; nail, finger-sucking, and stammering aids; eight wellness groups in A61H) and, in the post-acute design, treats continence and ostomy supplies (A61F5/44–49) as their own home-health unit. Pre-trend p is the joint test on each design's pooled event study; for FDA reclassification and competitive bidding, see the wild cluster bootstrap in Table 5. In the 1983 design the stricter list removes eight small wellness fields (220 patents) from the comparison group, which leaves the estimates unchanged and lowers the pre-trend p from 0.15 to 0.04. In the post-acute design it removes seven treated wellness groups in A61H, leaving 19 treated groups. Pre-trend tests for the FDA designs are reported in Table 5. +, *, **, and *** denote statistical significance at the 0.10, 0.05, 0.01, and 0.001 levels, respectively; for FDA reclassification III→II and competitive bidding, significance uses the wild cluster bootstrap.

Table A.13: Variance Decomposition of the Production-Cost Score

Source of variation in p^{cost}	**Share of total variance**
Between CPC-subclass	11.7%
Between filing-year	5.6%
Within a class × year cell (patent content)	**83.9%**

Note: Device production-cost score; $N = 401{,}888$ device patents with a filing year (13 CPC subclasses, filed 1976–2024). Most of the variance lies within class-by-year cells, in the content of individual patents. All specifications in the paper include year fixed effects.

Table A.14: Most Predictive Terms in Each Direction Classifier

Channel	Highest-weight terms (TF-IDF logistic classifier)
Main direction dimensions	
Production-cost-reducing	cost, manufacture, expensive, manufacturing, production, inexpensive, plastic, process
Quality-expanding	cells, bone, direction, activity, rate, disease, brain, cell
Episode-cost channels	
Length-of-stay-reducing	healing, incision, surgery, invasive, procedure, wound, patients, surgical
Complication-reducing	risk, infection, safety, damage, tube, medical, patient, contamination
Site-shifting	portable, invasive, user, device, ambulatory, require, home, pump
Other scored channel	
Labor-saving	time, time consuming, consuming, procedure, patient, requires, required, dental

Note: The eight highest-weight terms (unigrams and bigrams) in each deployed TF-IDF logistic classifier, ranked by coefficient. Length-of-stay-reducing, complication-reducing, and site-shifting aggregate into the episode-cost measure; labor-saving is not part of it.

Table A.15: Inter-Rater Agreement: Human, Frontier Models, and the Deployed Classifier

Channel	N	Human–human Cohen's κ	Human–classifier Cohen's κ	Human–Claude Cohen's κ	Claude–GPT-5 Cohen's κ
Cost	202	**0.44**	0.39	0.38	0.78
Quality	202	0.20	0.13	0.20	0.58
Complication	202	0.15	0.14	0.13	0.47

Note: Cohen's κ (labels binarized at 0.5) among two independent human coders, the deployed TF-IDF classifier, and frontier language models. The human–human column uses a blinded 150-patent subset double-coded by the author; the other columns use the full ∼200, coded by a research assistant. On the cost-reducing channel, the human coders agree at $\kappa = 0.44$ (75% agreement); on the same 150 patents, the classifier agrees with them at 0.45 and 0.66. On the full ∼200, the classifier and a frontier model agree with the single coder at 0.39 and 0.38. Agreement is lower on quality, where a near-universal base rate lowers κ despite 76% agreement, and on complications. Against the average of the two coders' ratings, the classifier's threshold-free AUCs are 0.84/0.78/0.76 (cost/quality/complication). The pairwise columns use the 202-patent validation set and are not comparable to the four-model range on the separate 643-patent sample ($\kappa = 0.40$–0.77) cited in the text.

Table A.16: Estimates under Four Classifiers

	TF-IDF (baseline)	PaECTER	GTE	DistilBERT
1983 inpatient PPS (1991–95)	3.31*	4.45***	2.25*	3.69***
	(1.31)	(1.22)	(0.89)	(1.02)
BPCI (2011)	1.58***	1.28***	1.15**	2.08**
	(0.46)	(0.31)	(0.35)	(0.80)
Post-acute PPS	1.09*	0.78	0.73	1.34*
	(0.48)	(0.58)	(0.52)	(0.57)
Pass-through expiry, production cost	1.02**	0.77	0.26	0.63
	(0.35)	(0.51)	(0.36)	(0.60)
Pass-through expiry, episode cost	−0.91*	−0.93*	0.02	−0.95
	(0.37)	(0.45)	(0.46)	(0.68)
Held-out AUC, episode-cost score	0.92	0.89	0.86	0.88

Note: Each column re-estimates the design in Table 4 with the channel scores of one classifier, on the patents scored by all four; estimates in percentage points, standard errors in parentheses. All four classifiers are logistic models fit to the same 5,460 language-model labels on the same text (title, abstract, and summary of invention). Post-acute: payment-rule outcome. Pass-through: exposed minus shielded years. Held-out AUC: 607 to 624 validation patents outside the training sample, against the training labeler, with sampling weights. +, *, **, and *** denote statistical significance at the 0.10, 0.05, 0.01, and 0.001 levels, respectively.

Table A.17: Distinctness of the Cost-Direction Measure

Direction score	Citations (impact)	Novelty (Kelly 2021)	Process/product (Bena–Simintzi)
Production-cost-reducing	−0.10	−0.05	−0.08
Quality-expanding	0.02	0.00	0.13
Complication-reducing	0.03	0.05	−0.04
Episode-cost-reducing (three channels)	0.09	0.04	0.04

Note. Pearson correlations of each cost-direction score with an external patent characteristic, after residualizing both on CPC main-group × issue-year fixed effects (5,707 cells); the text classifier does not use these characteristics. Citations are forward citations; novelty is − log backward similarity (Kelly et al., 2021); process/product is the majority-process-claims indicator of Bena and Simintzi (2025), the measure most adjacent to cost-versus-quality direction. Coverage runs 85–100% of the 411,651 device patents. Every correlation is at most 0.13 in absolute value—the largest, quality-expanding with process, is 0.13; the rest are at most 0.10—and the binarized agreement with the process classification is Cohen's $\kappa = -0.01$ for the production score and 0.04 for the episode score on these device patents. Correlations with Kogan et al. (2017) stock-market patent value (the 18.7% of patents at public firms) and the Kelly et al. (2021) breakthrough indicator are similarly small (all $|r| \leq 0.08$). Because the sample has roughly 400,000 patents, correlations this small can still differ statistically from zero, so the table reports magnitudes rather than significance. The measure captures the *direction* of invention, not its value, novelty, or process character.

Table A.18: Patent–HCPCS Crosswalk: Blinded Accuracy Audit

	Hand-coded ($n = 72$)		Language-model-coded ($n = 300$)	
Links	Exact	Same-domain	Exact	Same-domain
Control, high cosine	0.58	0.92	0.62	0.94
Treated, high cosine	0.75	0.92	0.50	0.90
All control	0.33	0.69	0.33	0.69
All treated	0.47	0.75	0.27	0.68

Note: Share of patent–HCPCS links that are true technological matches, on a stratified sample of 300 links (treated vs. control HCPCS in the main designs × three cosine tiers), coded blind to treatment status and cosine. "Exact" requires the specific HCPCS; "same-domain" accepts the correct clinical or technological family, the relevant standard for a category-level difference-in-differences design. The hand-coded 72-link subset and the language-model-coded 300-link pass agree at Cohen's $\kappa = 0.61$. Accuracy rises with cosine and does not differ detectably by treatment status.