How Does Technological Change Affect Quality-Adjusted Prices in Health Care? Systematic Evidence from Thousands of Innovations*

Kristopher J. Hult† Sonia Jaffe‡ Tomas J. Philipson§

June 15, 2017

Abstract

Medical innovations have improved treatment outcomes for many diseases but have simultaneously raised health care spending. Many health economists believe that technological change is the major factor driving the growth of the heath care sector. Whether quality has increased as much as spending – that is, whether new innovations raise or lower quality-adjusted prices in health care – is a central question for both positive and normative analysis of this sector. Previous research has only provided anecdotal evidence on this issue. We perform a systematic analysis of the impact of technological change on quality-adjusted prices, with over six thousand comparisons between innovations and incumbent technologies. For each innovation in our dataset, we observe its price and quality, as well as the price and quality of an incumbent technology treating the same disease. Our main finding is that for about two-thirds (68%) of innovations, the innovation's quality-adjusted price is higher than the incumbent's. Despite this finding, we argue that quality-adjusted prices may fall or rise over time depending on how fast prices decline for a given treatment. We calibrate that price declines of 4% between the time when a treatment is a new innovation and the time when it has become the incumbent would be sufficient to offset the observed price difference between innovators and incumbents for a majority of the innovations. We conclude by discussing the conditions particular to the health care industry that may result in less rapid declines, or even increases, in quality-adjusted prices over time.

*We are thankful to Casey Mulligan, session attendees at the University of Chicago, Tufts University, Bates White Life Sciences Conference, ASHEcon 2016, and participants at the BFI Health Sector and the Economy Conference for comments, and to the Becker Friedman Institute's Health Economics Initiative at The University of Chicago for financial support and the CEA Registry for data support. Ivy Sun, Deepon Bhaumik, and Kan Xu provided valuable research assistance.

†Charles River Associates. The conclusions set forth herein are based on independent research and publicly available material. The views expressed herein are the views and opinions of the author and do not reflect or represent the views of Charles River Associates or any of the organizations with which the author is affiliated. Any opinion expressed herein shall not amount to any form of guarantee that the author or Charles River Associates has determined or predicted future events or circumstances, and no such reliance may be inferred or implied. The author and Charles River Associates accept no duty of care or liability of any kind whatsoever to any party, and no responsibility for damages, if any, suffered by any party as a result of decisions made, or not made, or actions taken, or not taken, based on this paper. Detailed information about Charles River Associates, a registered trade name of CRA International, Inc., is available at www.crai.com.

‡Becker Friedman Institute, University of Chicago, spj@uchicago.edu. Becker Friedman Institute, Room 213; 1126 E 59th St; Chicago IL 60637.

§Becker Friedman Institute and Irving B. Harris Graduate School of Public Policy Studies, University of Chicago, tjphilip@uchicago.edu

1 Introduction

Given the rapid expansion of health care spending in most developed countries, there is an ongoing debate about whether this spending growth is accompanied by sufficient growth in the quality of care. Rapid medical innovation that enables us to treat previously untreatable diseases and improve existing treatments is a significant contributor to both spending growth and improvements in care. A central question of health economics is whether these innovations increase the value of care – that is, whether they raise the benefits of health care more than they increase the cost. In other words, do quality-adjusted prices rise or fall with new innovations?

In this paper, we provide new systematic evidence on how medical innovations affect quality-adjusted prices for thousands of innovations. We obtain such a large set of technologies by recognizing the economic content of medical "cost-effectiveness" studies, performed over the course of at least half a century. By using an economic lens to reinterpret the cost-effectiveness literature, we can gain more systematic insight into the impact of technological change on quality-adjusted prices than previous anecdotal findings for a few set of indicators. In particular, what this literature calls the "cost-effectiveness ratio" is the price of the technology divided by the quality of the treatment. It is a quality-adjusted price, analogous to the price per square foot of housing. The incremental cost-effectiveness ratio (ICER) compares the price and quality of a new treatment with a "comparator" – usually the incumbent technology representing the standard of care before the arrival of the new treatment. The goal of the ICER is to measure the marginal quality-adjusted price of the added quality provided by the new innovation. Therefore, studies reporting ICER levels often offer measures of both the quality and price of innovations and incumbent treatments. To provide an aggregate and systematic analysis, we use a database of the universe of cost-effectiveness studies from Tufts Medical Center called the Cost-Effectiveness Analysis Registry (CEAR), which contains over 4,000 cost-effectiveness studies. CEAR is an extensive data set of cost-effectiveness articles published in the peer-reviewed medical literature over the last 40 years.

Our main finding concerns the cross-sectional relationship between new innovations and incumbent technologies; that is, how the quality-adjusted prices between the two compare at a given point in time. Using the CEAR data, we find that the median innovation has a quality-adjusted price about 4 percent higher than the incumbent's quality-adjusted price, with 68 percent of new technologies having higher quality-adjusted prices than those of the incumbents. This estimate comes from a combination of slightly higher quality (median 1%) and moderately higher price (median 8%). We analyze potential drivers of differences between innovators and incumbents in quality-adjusted prices. We look at how an inno-

vation's price, relative to the incumbent's, varies with the type of treatment, the type of disease, and the difference in efficacy between the two treatments. We also consider regional differences potentially induced by third-party payment policies. We find that the innovator's quality-adjusted price, relative to the incumbent's, tends to be lower in the United States and European Union compared to other countries.

It is often argued that health care differs from other industries in that quality-adjusted prices rise over time instead of falling as they do in other industries, such as telecommunications. At first glance, our findings seem consistent with that pattern. However, this ignores the time series behavior of prices of a given technology due to competition. In particular, new therapeutic competition from related innovations and generic competition due to patent expiration often cause prices of a given treatment to fall over time. Therefore, the overall price of treating a disease may fall over time, even though new innovations have higher prices than incumbents in the cross section. In other words, a new innovation can be more expensive than the incumbent after entry and still be cheaper than the incumbent was before entry. We calibrate that if competition caused innovations to cut prices by at least 4.2% before the entry of subsequent new technologies, then overall quality-adjusted prices would fall for half of the markets, given the observed innovator premium in the cross-section.

This paper relates to several strands of work. Aggregate growth accounting attributes the residual from health care spending regression to the impact of technological change but does not measure innovation directly (Newhouse, 1992). Jena and Philipson (2008) use the CEAR data to address topics of a broader nature than the product-specific evaluations for which it was designed by analyzing the relationship between cost-effectiveness thresholds and innovation incentives. This paper also relates to existing anecdotal case studies that look at quality-adjusted price trends within a given indication, such as Cutler et al. (1998) for heart disease and Frank et al. (1999) for depression. We extend this literature by offering more aggregated and systematic evidence on these issues. Although there are some shortcomings in this systematic evidence, as we discuss in the conclusion, we believe using this evidence is superior to not using it and relying on anecdotal case studies. We value the breadth and volume of the CEAR data despite the limitations created by its aggregate nature, just as economists value the information available in national accounts while recognizing and working to improve their data shortcomings.

The paper proceeds as follows. Section 2 discusses the CEAR data on cost-effectiveness studies. Section 3 considers the cross-sectional findings of the relationship between price and quality of new innovations and incumbent technologies. Section 4 discusses time trends of quality-adjusted prices and potential explanations for the discussed findings. Finally, Section 5 summarizes the findings and conclusions.

2 CEAR Data

We use data from the Tufts Medical Center Cost-Effectiveness Analysis Registry (CEAR). CEAR is a dataset of the methods and findings of about 4,800 published cost-utility analysis articles some of which cover multiple treatments. The database is intended to be a comprehensive collection of all such articles published in the peer-reviewed medical literature. The registry uses two trained reviewers to independently review each article and collect a wide variety of variables. Their analyses are compared for accuracy.

Each article compares a newer treatment option to one or more standards of care (SOC). The registry collects and organizes data on article information, disease classification, study methodology, and measurement details. The data set covers articles published between 1976 and 2014 with the number of articles increasing from an average of one per year during the first ten years, to an average of about 475 articles per year during the last five years. As a result, 85% of the data comes from papers published between 2004 and 2014. About half of the articles, all written since 2002 have prices and qualities separately for the innovation and the standard of care. The other articles only have a single measure of the incremental cost efficiency ratio (ICER) – the difference in price divided by the difference in quality.[1]

Summary statistics for the main variables are in Table 1. Each article may provide multiple studies (comparing different subgroups, patient settings, or standards of care), so we have a total of 12,560 studies, about 6,500 of which have complete price and quality data. The main variables of interest are the ones related to cost and effectiveness. The price of an intervention (either the innovation or the standard of care) includes all the costs that the article was able to measure – both direct costs and non-health care costs.[2] Since the total cost is also the full price, we refer to it as the price of the treatment.[3] The price for both the new treatment and the SOC are measured per person treated. We convert them to 2014 US dollars using the medical CPI and yearly exchange rates.[4]

[1]ICER is commonly used in the cost-effectiveness literature as a measure of the marginal price for the additional quality gained from the innovation, relative to the standard of care.

[2]The data codebook defines: *Direct Medical Costs* as "Health care resource costs related to the intervention and its side effects. These costs include those impacts directly attributable to the intervention and those related to current and future consequences to the intervention (e.g. hospitalization, MD or other provider, long-term care, other health care which includes medications, outpatient procedures and laboratory costs)" and *Non-Health Care Costs* as "Non-health care [costs] resources related to the intervention and its side effects (e.g. travel time to doctor, caregiver time and workplace productivity impacts, transportation costs, patient productivity costs)." Only about a quarter of the studies list direct costs separately and less than 5% list non-health costs separately, so we focus on the total costs.

[3]About a fifth of the studies refer to "social cost" and try to include costs or savings from additional treatment needed or avoided, rather than just the price of the treatment. This more inclusive measure may be closer to the true price from a societal welfare perspective. We control for the type of cost measure used in our regressions, but it does not affect the results.

[4]We omit observations with negative cost for either the innovation or the SOC. We also omit observations

Table 1: Summary Statistics

	Mean	Min	Max	S.D.	Obs
Year	2010	1976	2014	5	12,560
US dummy	0.39	0	1	0.49	12,560
EU dummy	0.43	0	1	0.49	12,560
Measures Social Cost	0.20	0	1	0.40	12,560
Score of Study Reliability	4.69	1	7	1.00	12,557

	Median	5th	95th	S.D.	Obs
Innovator Quality	7.7	0.2	27.3	10.7	6,597
SOC Quality	7.3	0.1	27.0	10.7	6,572
Innovator Price	21,506	263	310,397	303,734	6,886
SOC Price	16,682	84	274,861	282,271	6,854
Innovator Price per QALY	4,532	36	88,345	162,054	6,504
SOC Price per QALY	3,755	20	83,505	196,667	6,391
ICER	17,415	-111,268	419,635	621,768	12,483

Note: This table summarizes the studies in the Cost-Effectiveness Analysis Registry. The US and EU dummies refer to the country in which the study was conducted. The 'Score of Study Reliability' is a rating that the reviewers compiling the database give to each study they read. SOC is the standard of care (incumbent) treatment. Quality is measured in quality-adjusted life years (QALYs). Prices are in 2014 US dollars. The 'ICER' is the incremental cost effectiveness ratio, which is the difference in prices between the innovator and the SOC divided by the difference in qualities.

Effectiveness is measured in quality-adjusted life years (QALY), which is a combination of the length of life and quality of life added by the treatment. A year of perfect health is equal to one QALY and a year of death is zero QALY, with different levels of health in between so that a person is indifferent between living x years at a QALY of $1/x$ and one year at perfect health. The effectiveness of a treatment is how many QALYs gives a patient, which we refer to as quality. Similar to costs, both the innovation and the SOC have quality measures.[5]

Prices of the innovations range dramatically, from around \$282 to \$315,324 between the 5th and 95th percentile since a wide variety of treatments are included. The median is around \$22,000. Quality also has a large range, between 0.26 and 27 QALY for the 5th and 95th percentile, with a median of 7.8 QALY. The treatments with higher price and quality include surgeries such as prophylactic oophorectomy, a surgery that reduces the risk of breast cancer and ovarian cancer, or treatments that have to be administered with very high frequency over a very long period, such as HIV antiretroviral therapies atazanavir-ritonavir or lopinavir-ritonavir.

One potential concern about using CEAR is that there may be heterogeneity in the quality of studies that are recorded by the registry. An unusual feature of these data is that each study in the registry is evaluated in terms of its quality through a scoring system. In compiling the dataset, the readers of the registry rate each article on a scale from 1 to 7, based on perceived correctness and comprehensiveness. This feature tells us which observations we should rely on more. We report our findings for the overall sample of studies as well as for the *high-score studies* – studies that have the median score rating or higher.

In discussing the other study characteristics, we limit the sample to the 6,472 studies for which we have price and quality data for both the innovation and the standard of care. There are studies from 70 countries, but North America and Europe account for 81% of the studies, including 38% from the United States and 17% from the United Kingdom. The dataset includes variables on the primary disease addressed by the treatment, the treatment's type of intervention, and the study sponsor. There are 65 disease categories, with about 50% of the studies coming from the five most studied diseases (infectious diseases (12%), cardiovascular diseases (12%), malignant neoplasms (12%), musculoskeletal and rheumatologic diseases (8%), neuro-psychiatric/neurological conditions (4%)) Each study covers one or more types of interventions including pharmaceutical (54%), surgical (11%), screening (18%) and medical procedures (11%).[6] The studies include sponsors with the main sponsors being governments

where the ICER, price, or price per QALY for the innovation or the SOC is over ten million dollars.

[5] We omit observations with quality values greater than 100 since it does not make sense for a treatment to add more than 100 years to someone's life. We also omit studies with negative quality values.

[6] The intervention types are *Care Delivery:* Provision of care; development of facility or distribution of

(38%), industry pharmaceutical, biotech, and medical device companies (32%), and non-profit organizations (8%). (As with interventions, there may be multiple sponsors for a treatment, so the sponsor indicators average to greater than one.)

There may be selection issues in determining which articles get studied and published, which would bias our attempts to understand how the average new innovation affects quality-adjusted prices. For instance, if only the most cost-effective new treatments get studied or there is a publication bias in favor of findings of high cost-effectiveness, that would bias our estimates. For selection to be an issue, it would need to occur independent of the intervention type, disease, country, and sponsor type since we control for these. We do not have data on market share, so we are comparing product level differences in price and quality between the innovation and the standard of care.

3 Comparing Innovators and Incumbents in the Cross-section

The measurements provided by the CEAR are highly relevant to economists interested in the value of health care in terms of quality-adjusted prices for health. Figure 1 below depicts a schematic of the variables measured in the CEAR in a traditional quality-price space. The incumbent technology corresponds to the standard of care and has quality and price (q_s, p_s), while the new innovation has quality and price (q, p). The slopes from the origin represent the quality-adjusted prices; the slope between the incumbent's and the innovator's quality-price pairs represents the ICER. The figure illustrates that the average price of quality (quality-adjusted price) of the innovator is larger than that of the incumbent when the marginal price (ICER) is higher than the quality-adjusted price of the incumbent. Put simply, the average rises if and only if the marginal price is higher than the previous average.

If one imagines centering a graph around the incumbents price and quality, (q_s, p_s), the quadrants represent combinations of price and quality differences that the new innovator can represent. For example in the northeast quadrant the innovator has both higher quality

personnel (e.g. a policy that changes the nurse-to-patient ratio, patient self-management program). *Health Education or Behavior:* An intervention designed to educate individuals on behaviors that promote, maintain or restore health (e.g. smoking cessation and prevention program). *Pharmaceutical:* Any drug or biotech product used for medical treatment or prevention (e.g. Lovastatin, Herceptin). *Surgical:* Invasive; cutting involved (e.g. appendectomy, transplantation – although bone marrow transplantation would be a medical procedure). *Immunization:* Receipt of vaccination (e.g. flu vaccine, HPV vaccine). *Diagnostic:* A method used to determine if and what type of disease is present (e.g. imaging, biopsy, PET scan, x-rays, in-vitro testing). *Medical Procedure:* Non-surgical, non-diagnostic procedures (e.g. angiogram, blood donation, mole removal, casting). *Medical Device:* May or may not require a surgical or implantation procedure (e.g. pacemaker, insulin pump, leg brace, and crutches). *Screening:* Refers to measures that detect disease (or test for risk factors) before it is symptomatic (e.g. breast cancer screening mammogram). *Other:* Any intervention not described above (e.g. injury prevention, food safety, or environmental health).

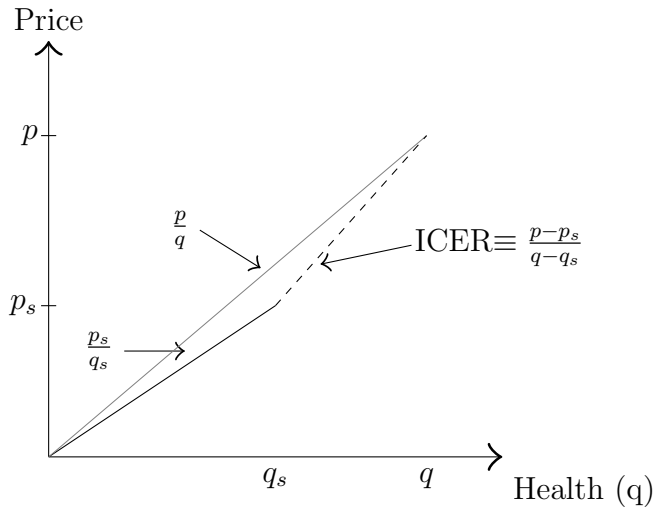


Figure 1: Price-Quality Space.

and price; within that quadrant, quality-adjusted prices rise if price differences dominate quality differences. Figure 2 provides the unconditional scatter plot of the joint distribution of price and quality differences between innovations and incumbents. In this figure, the origin represents an incumbent's quality and price, and the axes reflect the percentage difference in quality and price between the innovator and the incumbent. The dashed 45° line represents when the innovator has the same quality-adjusted price as the incumbent which corresponds to when the percentage difference in price equals the percentage difference in quality: $\frac{p}{q} = \frac{p_s}{q_s}$ implies $\frac{p-p_s}{p_s} = \frac{q-q_s}{q_s}$.

In general, both price and quality are higher for the new innovation relative to the incumbent. There is a price increase in 78% of observations and a quality increase in 85%. The distribution is *very* skewed to the right. The average innovation increases price by 139% relative to the incumbent, but the median is an 8% increase; the average quality difference is 26%, but the median is 1%. The percentage change in price has a much wider range than the percentage change in quality. The 31% of innovations below the 45° line (green) in Figure 2 have lower quality-adjusted prices. The innovations with higher quality-adjusted prices than the incumbent are split into the 11% that lower quality (red) and the 56% that improve quality (blue). Even though these innovations have higher quality-adjusted prices, they can consumers better off relative to incumbent treatmentss if they increase quality.

With extreme substitutability of demand across treatments, we would not expect to see treatments with a higher price and lower quality or a lower price and higher quality (the off-diagonal regions in Figure 2 where $(p_2 - p_1) \cdot (q_2 - q_1) < 0$). This is broadly true in the data, as 69% of the innovations lie in the diagonal quadrants where higher quality treatments

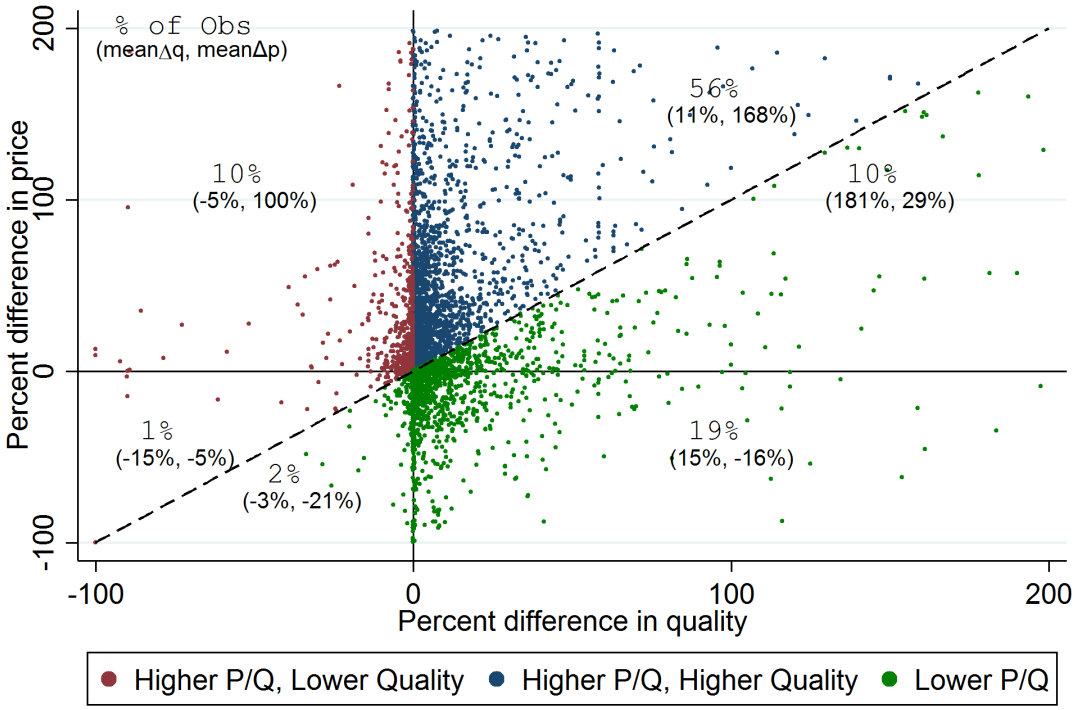


Figure 2: Price and quality differences.

Note: The x-axis is the difference in quality between the innovation and the incumbent as a percent of the incumbent's quality. The y-axis is the same measure for price. The top number in each section is the percent of observations that fall in that quadrant or octant, including the outliers that are censored from the graph. They do not sum to 100% because of the 2% of observations that fall on one of the axes. The numbers in parentheses are the average percent difference in quality and price for observations in that quadrant.

command higher prices. However, note that there is a non-trivial fraction of innovations with higher quality and lower price (19%), but nevertheless do not destroy the market for the SOC. There are also some innovations with that enter with a higher price and lower quality (10%). The existence of these innovations may be due to heterogeneity in treatment effects. Another explanation is that the first type of innovation will eventually replace the SOC after being shown to be of higher value and the latter will exit after it becomes evident they are of lower value.

The median quality-adjusted prices for new innovations and the SOC, measured as the cost per QALY, are \$2,700 and \$2,400 respectively. Figure 3 plots the distribution of the ratio of the quality-adjusted prices of the new innovation and the SOC, $\frac{p}{q}/\frac{p_s}{q_s}$. If this *quality-adjusted price ratio* is above one, it means the new innovation had a higher quality-adjusted price than the SOC. The histogram indicates that a high fraction of new technologies do not change quality-adjusted prices much, but new innovations do tend to have higher quality-adjusted prices than the SOCs. The median ratio is 1.04, meaning the quality-adjusted price of the innovation is 4% higher than the standard of care; the distribution is very right-skewed with a standard deviation of 32, so only 2.4% of innovations have cost effectiveness ratios that are statistically larger than one. If we focus on highly scored studies, the median is about the same, but the standard deviation is even higher. Only 0.9% of those studies are statistically different from one.

These are the overall effects for all disease categories and modes of intervention. Table 2 breaks down the change in quality-adjusted price for the most common disease categories. Infectious disease, malignant neoplasms, and breast cancer have the highest quality-adjusted prices relative to the incumbents. Table 3 does the same for intervention types.[7] Pharmaceuticals are the intervention type with the highest relative price. Educational interventions tend to be less expensive relative to the incumbent and also vary less. The differences across intervention type are not as robust to the exclusion of low-scored studies as the disease categories. These tables indicate that the large variation we see in the relative cost effectiveness of an innovation is not due primarily to differences across disease categories or treatment types. There is more variation within categories than across categories.

One reason innovators may command high quality-adjusted prices is if they offer "breakthrough" innovations. The large treatment improvement from these innovations may allow the innovator to charge a higher price. Since the quality is also higher, the effect on quality-adjusted price is ambiguous. To investigate these effects we examine the impact of the quality difference $(q - q_s)$ on the relative price. As shown in Table 4, for highly-scored

[7]Since studies may have multiple intervention types, the sum of the number of observations is greater than the number of studies.

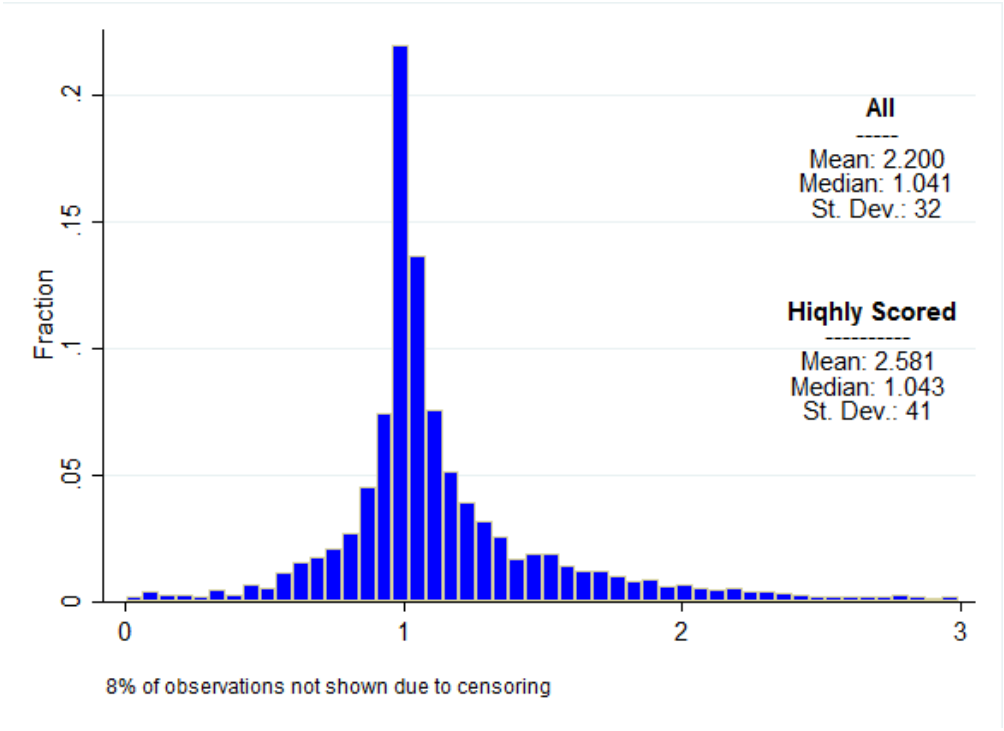


Figure 3: Ratio of quality-adjusted prices.

Note: This graph shows the distribution across innovations of the ratio of innovator's and incumbent's quality-adjusted prices. For readability, the graph omits studies with a cost effectiveness ratio above 3, thereby censoring 8% of observations. 'Highly Scored' refers to studies with at least the median rating given by the readers.

Table 2: Percent difference in quality-adjusted price by disease type

	All				High Score			
	Obs	Median	Mean	St. Dev.	Obs	Median	Mean	St. Dev.
Maternal/Child	65	-0.17	-7.60	37.73	23	-0.17	-9.60	39.74
Genito-Urinary	277	0.25	17.44	75.21	148	0.55	14.56	53.85
Neuro-Psychiatric/ Neurological	490	1.57	21.05	104.59	311	1.55	13.33	74.61
Endocrine Disorders	413	1.71	22.14	118.30	234	3.11	22.09	98.59
Cardiovascular	1176	2.87	28.48	98.04	796	2.76	21.57	72.67
Musculoskeletal/ Rheumatologic	633	3.00	31.25	124.46	356	3.55	36.95	135.27
Digestive	300	6.37	44.25	116.14	136	0.99	31.81	93.37
Respiratory	158	7.41	42.86	159.29	120	7.47	42.12	152.17
Malignant Neoplasms	1064	7.58	42.37	122.31	637	7.49	47.70	137.01
Sense Organ	108	8.19	57.50	125.11	62	11.59	72.02	127.88
Infectious	973	8.91	42.26	126.71	626	9.87	47.03	134.39

Note: This table shows the percent difference in quality-adjusted prices, $100 \cdot \left(\frac{p}{q}/\frac{p_s}{q_s} - 1\right)$, across innovations by disease category. It omits the top 1% of quality-adjusted price ratios. The right panel shows only studies with at least the median score – the rating of the study given by the readers.

Table 3: Percent difference in quality-adjusted price by intervention type

	All				High Score			
	Obs	Median	Mean	St. Dev.	Obs	Median	Mean	St. Dev.
Education	319	0.78	13.70	79.69	217	2.14	16.15	59.30
Diagnostic	528	1.35	19.16	88.48	297	1.55	27.89	106.68
Care Delivery	498	1.49	28.18	99.09	255	1.71	20.03	70.00
Screening	1109	2.59	31.57	122.83	676	3.17	43.72	149.00
Surgical	697	3.43	25.54	93.79	294	9.77	37.09	100.33
Immunization	193	3.99	64.37	163.49	128	2.25	41.05	138.41
Device	491	5.23	29.22	94.56	250	7.76	26.80	57.56
Pharmaceutical	3448	5.67	40.36	128.43	2220	4.93	37.80	120.62
Procedure	696	6.38	34.24	100.19	351	7.49	38.82	102.23

Note: This table shows the percent difference in quality-adjusted prices, $100 \cdot \left(\frac{p}{q}/\frac{p_s}{q_s} - 1\right)$, across innovations by intervention type. Studies may have multiple intervention types, so the sum of the number of observations is greater than the number of studies. The table omits the top 1% of quality-adjusted price ratios. The right panel shows only studies with at least the median score – the rating of the study given by the readers.

studies, innovations in the top quartile of quality improvement tend to have higher quality-adjusted price ratios. The mean is about 0.15 higher, and the median is 0.034-0.06 higher, meaning the percent difference in quality-adjusted price is 3-6 percentage points higher. The difference is much smaller and statistically insignificant if we look at all studies instead of just the highly-scored ones, so it is not clear how robust this relationship is. Note that the last column of Table 4 controls for the type of intervention, disease, and sponsor. If certain types of interventions or diseases have higher quality-adjusted price ratios because they have higher quality improvements, we would want to control for this effect.

Table 4: Quality-adjusted price ratio by size of quality difference

			Quantile Regression	
	(1)	(2)	(3)	(4)
50th-75th Percentile Quality Change	0.0320	0.110*	0.00749	0.000457
	(0.0515)	(0.0547)	(0.0107)	(0.0142)
> 75th Percentile Quality Change	0.159**	0.197***	0.0613***	0.0346*
	(0.0515)	(0.0570)	(0.0107)	(0.0148)
Controls:				
Year Dummies	X	X	X	X
Intervention, Disease, Cost, & Sponsor Type		X		X
Observations	3783	3783	3821	3821
Adjusted R^2	.01114	.04947	.0013	.01038

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table shows coefficients from regressions of the quality-adjusted price ratio, $\frac{p}{q}/\frac{p_s}{q_s}$, on dummies for the top two quartiles of quality change, $\frac{q-q_s}{q_s}$. The first two columns are linear regressions and omit the top 1% of quality-adjusted price ratios; the last two columns are median regressions. Columns (2) and (4) include dummies for each type of intervention, disease category treated, the type of sponsor the study had and the type of cost it measured. All regressions include only highly-scored studies – those which received at least the median ranking from the readers.

Another factor that may drive price differences are the price controls more frequently imposed in Europe than in the United States. Price controls may affect both the new innovation and incumbent, so they may not affect the quality-adjusted price ratio. As shown in Table 5, there is no systematic difference in the ratio of quality-adjusted price between studies done in the United States and in Europe, including the United Kingdom. However, these countries and Japan have substantially lower quality-adjusted price ratios than studies done in the rest of the world. In the US and EU, the quality-adjusted price ratio is about 0.2-0.3 lower on average; the median is about 0.06 lower. These results are qualitatively

unchanged if we look at all studies instead of just highly rated ones.

Table 5: Quality-adjusted price ratio by geography

			Quantile Regression	
US dummy	-0.353***	-0.302***	-0.0881***	-0.0665***
	(0.0622)	(0.0653)	(0.0128)	(0.0169)
EU dummy	-0.383***	-0.233***	-0.0994***	-0.0627***
	(0.0639)	(0.0683)	(0.0132)	(0.0177)
UK dummy	0.0453	0.0192	0.0262*	0.0189
	(0.0632)	(0.0675)	(0.0131)	(0.0175)
Japan	-0.399**	-0.367*	-0.0525	0.00168
	(0.155)	(0.158)	(0.0321)	(0.0412)
Constant	1.534***	1.408	1.126***	1.329***
	(0.143)	(0.762)	(0.0295)	(0.199)
Controls:				
Year Dummies	X	X	X	X
Intervention, Disease, Cost, & Sponsor Type		X		X
Observations	3783	3783	3821	3821
Adjusted R^2	.01936	.05171	.00161	.01079

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table shows coefficients from regressions of the quality-adjusted price ratio, $\frac{p}{q}/\frac{p_s}{q_s}$, on dummies for the country or region in which the study took place. The first two columns are linear regressions and omit the top 1% of quality-adjusted price ratios; the last two columns are median regressions. Columns (2) and (4) include dummies for each type of intervention, disease category being treated, the type of sponsor the study had and the type of cost it measured. All regressions include only highly-scored studies – those which received at least the median ranking from the readers. Studies in the UK are also included in the EU category, so the coefficient on the UK dummy is relative to the level in the EU.

Pharmaceuticals make up about half of our sample, and the high prices of new drugs have received a lot of affections in both academic and policy analysis. Table 6 looks at how the quality-adjusted price ratios for pharmaceutical innovations in highly-rated studies differ in different time blocks. Pharmaceutical innovations have somewhat higher prices than others, with the effect driven by studies from 2012-2014 having a median (mean) quality-adjusted price ratio that is 0.058 (0.145) higher than other innovations.[8] These results are consistent

[8] If we look at all studies, instead of just highly rated ones, the coefficient for 2012-2014 is slightly larger and the median in quality-adjusted price ratio for pharmaceuticals is also somewhat larger than other innovations in 2007–2010.

with the idea that new drugs have been more expensive to an extent not accounted for by their therapeutic value.

Table 6: Quality-adjusted price ratio of pharmaceuticals

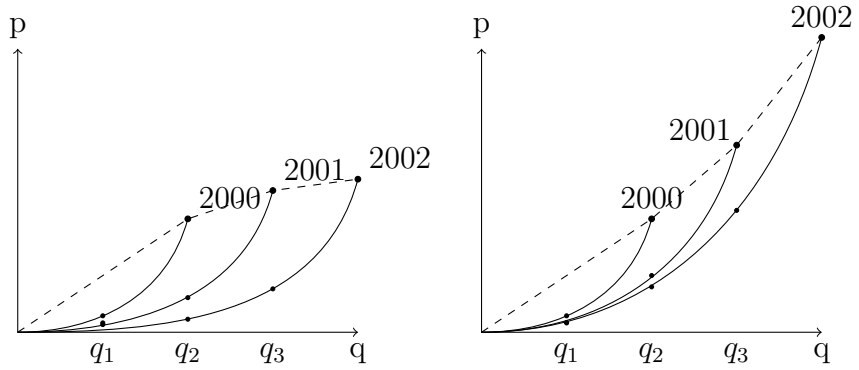
			Quantile Regression		
	(1)	(2)	(3)	(4)	(5)
Pharmaceutical	0.0428		0.0167*		
	(0.0427)		(0.00715)		
Pharma and 2002-2006		0.128		-0.0326	-0.00356
		(0.131)		(0.0237)	(0.0343)
Pharma and 2007-2010		-0.160		-0.0473**	-0.0152
		(0.0861)		(0.0156)	(0.0229)
Pharma and 2011-2014		0.108*		0.0383***	0.0638***
		(0.0532)		(0.00969)	(0.0138)
2006-2011		0.226		0.0104	-0.000734
		(0.123)		(0.0223)	(0.0327)
2012-2014		0.0348		-0.0482*	-0.0577
		(0.112)		(0.0203)	(0.0312)
Controls:					
Disease, Cost, & Sponsor Type					X
Observations	3783	3783	3821	3821	3821
Adjusted R^2	0	.00111	.00011	.00058	.00905

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table shows coefficients from regressions with the quality-adjusted price ratio, $\frac{p}{q}/\frac{p_s}{q_s}$, as the dependent variable. It uses only highly-scored studies – those which received at least the median ranking from the readers. The first two columns are linear regressions and omit the top 1% of quality-adjusted price ratios; the last three columns are median regressions. Columns (1) and (3) look at the difference between pharmaceutical and non-pharmaceutical innovations; Columns (2), (4), and (5) interact the pharmaceutical dummy with three time periods. Columns (5) includes dummies for each disease category treated, the type of sponsor the study had and the type of cost it measured.

4 Implications for the Time Series of Prices

Our evidence suggests that the quality-adjusted price of an innovation is often higher than the incumbent standard of care once the new innovation has entered the market. However, the overall time trend in prices does not depend only on the cross-sectional difference between innovators and incumbents, but also on the difference in price over time within products.

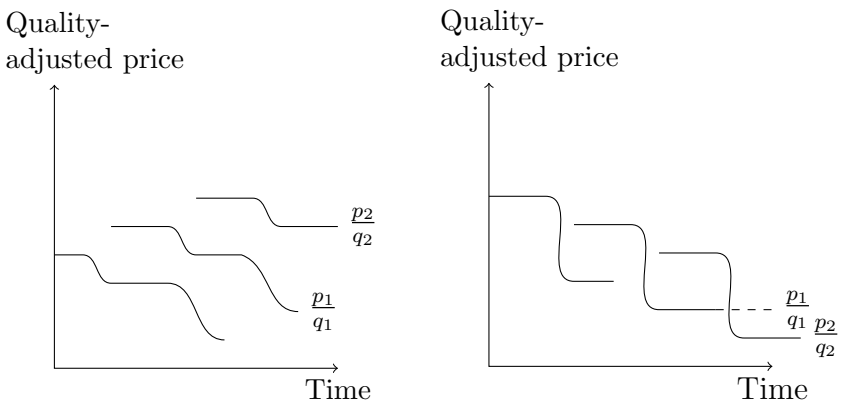


(a) Decreasing $\frac{p}{q}$ (b) Increasing $\frac{p}{q}$

Figure 4: Different possible price trends.

Figure 4 shows two hypothetical cross sections of price and quality relationships over time where each subsequent entry raises quality, as in a quality ladder model. Each year, a new product enters with a higher quality and the prices of the incumbents fall. In both cases, the price-quality relationship is convex within a year: the entrant has a higher quality-adjusted price than existing products, so quality-adjusted price is increasing in quality. However, the fact that new innovations may have higher quality-adjusted prices in the cross section does not restrict the path of quality-adjusted prices of new innovations over time – the top point on each line. The overall trend in quality-adjusted price can be increasing or decreasing. Figure 5 demonstrates this same idea by considering the quality-adjusted price of each product over time. Figure 5a shows a market where the price of each product falls over time, particularly when a new entry occurs, but the higher prices of new products cause an overall upward trend in quality-adjusted prices. Figure 5b shows a market where the fall in prices upon entry is large relative to the cross-sectional difference between the incumbent and the entrant, so there is an overall downward trend. Innovators having a higher price at a given point in time is consistent with both, for example, Cutler et al. (1998) who found falling quality-adjusted prices over time for heart attack treatments, and Howard et al. (2015), who found rising quality-adjusted prices over time for oncology drugs.

More precisely, let us think of the standard of care, with pre-entry price and quality p_s and q_s, as a composite of all the goods in the market prior to the innovation's entry. Then the pre-entry, quality-adjusted price was just p_s/q_s. The innovation enters with quality q, post entry prices are p'_s and p', and the incumbent's market share is ms'. The change in the



(a) Pre-entry Trends Leading to Overall Upward Trend in Prices.

(b) Pre-entry Trends Leading to Overall Downward Trend in Prices.

Figure 5: Different possible pre-entry trends in the price for the standard of care.

market average quality-adjusted price (QAP) is

$$QAP' - QAP = \left(ms' \frac{p'}{q} + (1 - ms') \frac{p'_s}{q_s} \right) - \frac{p_s}{q_s} = ms' \left(\frac{p'}{q} - \frac{p'_s}{q_s} \right) - \frac{p_s - p'_s}{q_s}.$$

Even if $\frac{p'}{q} > \frac{p'_s}{q_s}$, quality-adjusted prices fall over time if the pre-entry incumbent price was sufficiently high:

$$p_s > ms' p' \frac{q_s}{q} + (1 - ms' p'_s) \equiv p^*_s.$$

If the incumbent's price is relatively unchanged between when it entered and when the innovation enters, then even a small difference in the quality-adjusted prices between the innovator and the incumbent could indicate an overall upward trend in prices. Conversely, if the incumbent's price has dropped dramatically since initial entry (for example, if the incumbent is a drug that is available in a generic form), it is likely that even with a substantial price difference between innovator and incumbent, there may be an overall downward price trend in the market. Even if the incumbent is a relatively recent innovation, the entrance of the innovator may itself generate competition that causes the incumbent to drop its price. Medical innovation may reduce prices even though quality-adjusted prices of new entrants are higher than those of the incumbents in the cross section.

It is beyond the scope of this paper to estimate the price drops and market shares. Instead we calculate for a given market share, $ms = \{60\%, 80\%, 100\%\}$ and a given price drop, $x = \frac{p^*_s - p_s}{p^*_s} \cdot 100$, which innovations would have market average QAP that were lower post-entry than pre-entry. Figure 6 shows the percent of innovations for whom the market

average QAP decreases with entry for different market shares and price drops. With zero price drop, market average quality-adjusted prices are lower for 32% of innovations, since they have lower quality-adjusted prices than the incumbent post-entry. For the other 68% of innovations, if they gain 100% market share, the median innovation would have no change in quality-adjusted prices with a 15% drop in the incumbent's price. If the incumbent prices dropped 75%, then 95% of innovations would have prices lower than the incumbent's pre-entry price. If the new innovation only gets 60% market share, then a 10% drop in the incumbent's price would mean the market average quality-adjusted price was unchanged at the median (of those innovations with higher quality-adjusted prices than the incumbent); a price drop of only 64% is needed for 95% of all innovations to have prices lower than the incumbent's pre-entry price.

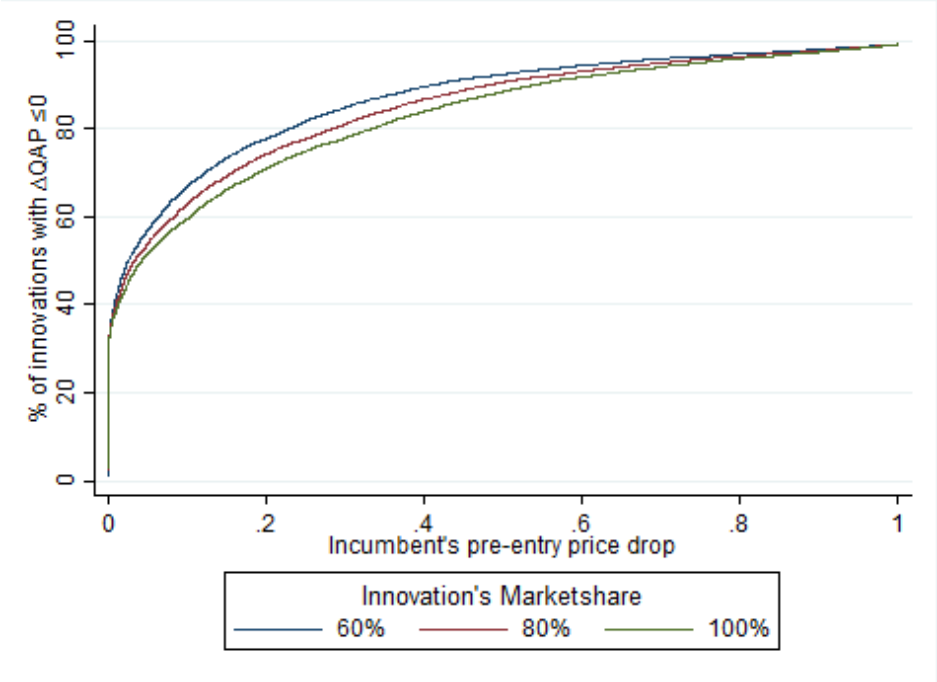


Figure 6: CDF of offsetting price drop

Note: This graph shows the cumulative distribution across innovations of incumbent's pre-entry price drop necessary to generate no change in overall price levels, for different entrant market shares. For each innovation and market share, we calculate the average quality-adjusted price in the market and then the incumbent price p_s^* that would give that average prior to the innovation's entry. For a given difference between the incumbent's pre-entry and post-entry prices (as a fraction of the pre-entry price) on the x-axis, the y-axis gives the percent of innovations for which the average market quality-adjusted price would be lower after the innovation.

4.1 Illustration for two important case studies

Much of the policy discussion of pharmaceutical pricing in the last few years is centered on new 'specialty drugs' and whether their incremental benefits justify their higher prices,

i.e. whether they have higher or lower quality-adjusted prices. We use two specialty drug classes: treatments for the hepatitis C virus (HCV) and multiple sclerosis (MS) to illustrate potential differences in time series and cross-sectional pricing patterns. These two drug classes have received much policy attention, the former for providing a breakthrough innovation through a cure and the latter because of rapidly rising prices of the same product over time. These cases are used to illustrate Figure 5, which illustrates the theoretical possibility of prices falling or rising over time, depending on the higher price of innovators in the cross-section is offset by decreases in incumbent prices over time.

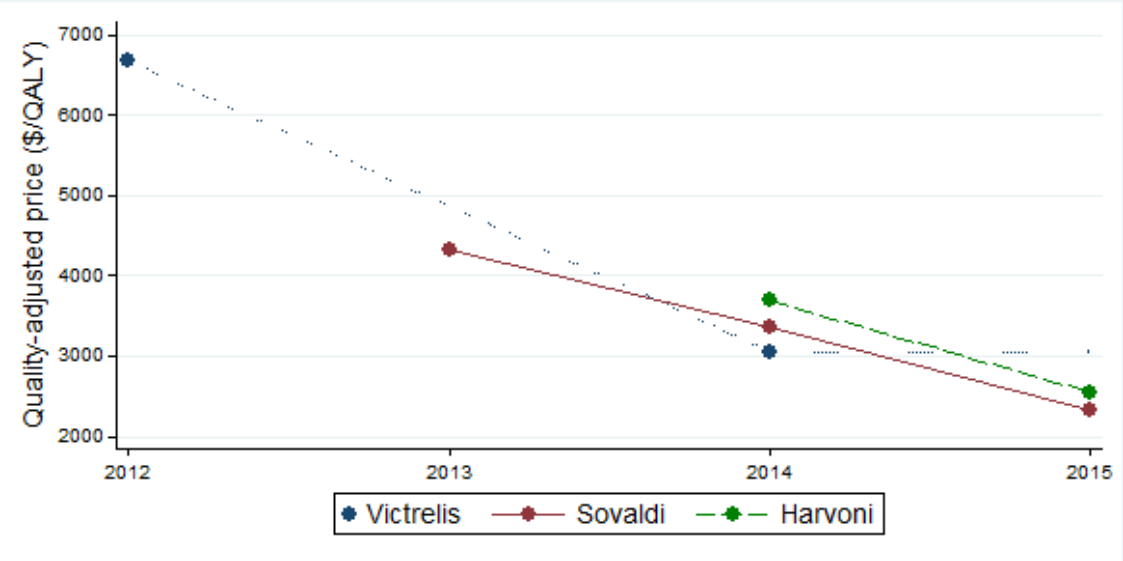


Figure 7: Hepatitis C drugs

Note: This graph shows the quality-adjusted prices for three Hepatitis C drugs over time. For Victrelis, the data are from two studies in the CEAR. For Sovaldi the prices come from Beasley (2015) and the QALY is from a 2013 CEAR study. For Harvoni, the QALY and initial price are from Zhang et al. (2015) and the 2014 and 2015 price discounts are from Beasley (2015).

In the case of hepatitis C, Figure 7 shows the price patterns for the different products over time, with the entry date defined as the start of the series. Because the HCV drugs are fairly new, there are relatively few observations in the CEAR data, so we supplement from other sources. For Victrelis, CEAR has price and QALY estimates in 2012 and 2014, from which we calculate quality-adjusted prices. For Sovaldi we use the company's prices and announced discounts for 2014 and 2015 (Beasley, 2015), combined with the QALY estimate from a 2013 CEAR study. We get the QALY and initial price for Harvoni from Zhang et al. (2015), which is the same type of study as the articles in CEAR, just published more recently. We again use the announced discounts to adjust the price over time.[9]

[9]For example, the 2,552 $/QALY number for Harvoni in 2015 is the $94,500 that Zhang et al. (2015) say

Sovaldi was the first innovation which offered a major increase in quality through essentially curing HCV through a 3 month treatment. As the graph indicates, the incumbent Victrelis dropped its price dramatically around Solvaldi's entry. Thus even though comparing their quality-adjusted prices in 2014, just after this entry occurred, would show the new innovation was more expensive, there was a substantial decrease in quality-adjusted prices. The HCV case is illustrative of the more general idea that when therapeutic price competition occurs through innovation, cross sectional and time series differences can be offsetting.

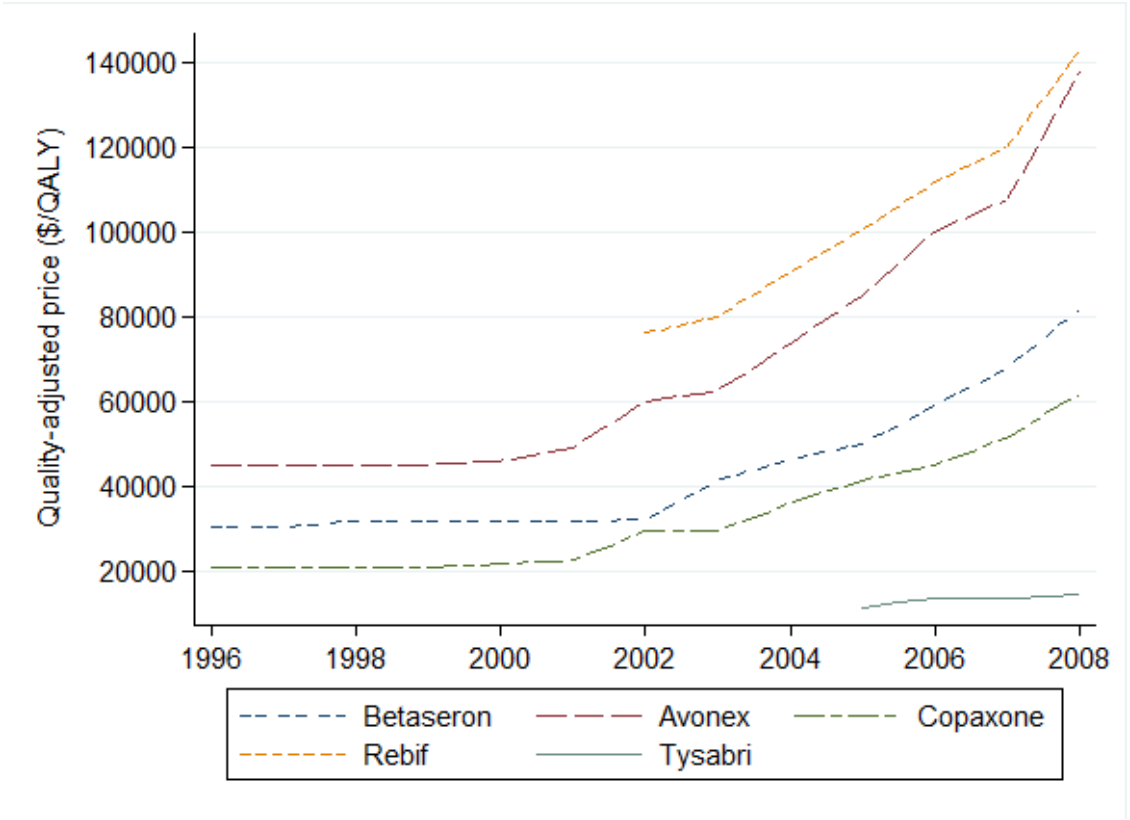


Figure 8: Multiple sclerosis drugs

Note: This graph shows the quality-adjusted prices for five multiple sclerosis drugs over time. The QALY estimates are the average from the CEAR studies and yearly prices are from Hartung et al. (2015).

The price patterns for multiple sclerosis drugs show a very different story. Figure 8 shows quality-adjusted prices for MS drugs over time. These drugs have been around longer, so there are more observations in CEAR, but none of the drugs appear yearly. We use the (average) QALY estimate from the CEAR and yearly prices from Hartung et al. (2015) to get yearly quality-adjusted prices. If anything, the entry of Rebif in 2002 coincides with the

a course of treatment costs, divided by their estimated quality of 20 QALYs times .54 to account for the 46% discount reported by Beasley (2015).

incumbents raising prices and the prices for all drugs are increasing overtime. Therefore, the difference in quality-adjusted price in 2004 between the new innovation, Rebif, and the incumbent, Avonex, understates the upward trend in prices in the market.

New drugs are less likely to cause price decreases when they are not good substitutes for the incumbents. There are three reasons that MS treatments are less substitutable than treatments for some other diseases. First, there is enormous heterogeneity in the MS population and treatment response, so chemically-similar drugs may have very different outcomes for the same patient (Lucchinetti et al., 2000). Second, MS is managed and not cured by treatments, so patients often need to take a range of MS drugs over their course of therapy. Third, there are significant differences across MS drugs in their efficacy, tolerability, and mode of administration. For example, Avonex is an injection treatment that is often used a first line treatment because it has mild side effects though low efficacy. Tysabri, which is administered by infusion, is often used as a later line treatment because it has high efficacy and potentially serious side effects (Smith et al., 2010). Since the drugs are weaker substitutes, they do not compete as much, so new drugs are less likely to cause incumbents to lower their prices. The MS case is illustrative of the more general idea that absent therapeutic price competition, prices for a given product may increase over time, so the cross sectional and time series differences may be mutually reinforcing.

5 Conclusion

We compare quality-adjusted prices of innovators and incumbents by providing an economic interpretation of the medical cost-effectiveness literature. We find that two-thirds of innovators have a higher quality-adjusted price than those of the incumbents; the median difference in quality-adjusted price is 4%. We find no systematic differences by disease type, though pharmaceuticals have somewhat larger differences, especially in recent years. There is some evidence that an innovator's quality-adjusted price is particularly likely to be high relative to that of the incumbent when there is a large difference in quality between the two. Studies conducted in the European Union and the United States have somewhat smaller quality-adjusted price ratios than studies performed elsewhere, but innovations still tend to be more expensive than the incumbent technologies. We show how quality-adjusted prices may still fall over time, depending on the price trends of the incumbents: a price decline of 4% between a product's entry and the entry of the subsequent innovation would be sufficient to generate a net decrease in quality-adjusted prices over time in the majority of products.

The data we use has several shortcomings that a more complete collection of CEA studies with a larger set of measures could potentially overcome. One issue is the selection of technologies studied, which may not be representative of all technologies. In particular,

CEA studies may be done more often for innovations that are costlier or higher quality. Second, CEA studies may vary in their perspective and therefore contain different costs (though if this is true for both the standard of care and the new innovation, these costs are measured uniformly across the two). Third, there may be several CEA studies on the same technology, in which case conflicting results may be obtained. However, we believe some of these shortcomings apply even more so to the reliance on anecdotal case studies to analyze the impact of technological change.

Overall, we view the study of CEAR-type data as a useful and evolving process in which improving the data over time is feasible and would lead to further improved analysis. These improvements are not much different from the improvements of national economic accounts made to address their shortcomings over time. We do not want to dismiss the informative value of national accounts while they are being improved. Likewise, we believe that relying on case studies rather than these more systematic data would be misguided. The question is not whether the data are perfect, but rather how they compare to alternative data for studying this important topic.

We end by discussing some reasons why innovation may affect quality-adjusted prices differently in health care than other industries, such as telecommunications, where next generation technologies often reduce quality-adjusted prices. Health care innovations may have fewer substitutes than those in other industries: a "first in class" designation can give a firm substantial market power, which may persist longer for innovations in health care than innovations in other industries because the FDA approval process can delay new entrants. Even average levels of market power may enable innovators to capture increased consumer value for treatments with a higher price; in this case, unmeasured cost offsets and complementarity between a treatment and either income or the prevailing level of health can lead to price increases over time.

If a new innovation is only one component of the full set of costs associated with a diagnosis, then its impact on the other episode costs – "cost offsets" – may be an important part of the value of the innovation. Drugs are usually a small fraction of the total episode cost but may generate cost offsets by preventing future doctor or hospital visits. If monopoly power allows the entrant to capture these cost offsets in a higher price for the innovation, then the quality-adjusted price may seem higher if the cost savings are not measured (though true total costs per quality may be lower). This explanation for higher quality-adjusted prices of innovators could be tested empirically if total costs and treatment costs could be separated, in which case cost offsets should have a positive effect on the innovator's premium.[10]

[10]The price variable that we use is the most inclusive one available, but non-health costs were recorded for only a fraction of the studies; they may not have captured all indirect costs.

Two forms of health-related complementarities may raise quality-adjusted prices over time. One is the complementarity between health and more health care: a greater level of health over time raises the value of additional health going forward. For example, the value of treating a life-threatening disease is higher the longer you live in the absence of the disease. As discussed by Dow et al. (1999), such complementarities are implied by competing risks models of mortality, which essentially involve a Leontief production function of overall length of life from competing cause-specific lifetimes. If a healthier population places a higher value on health improvements from a given disease and this can be priced out by new innovators, quality-adjusted prices may rise with the baseline health level.

The second health complementarity is between health and consumption (Hall and Jones, 2007). The willingness to pay for longevity increases with economic growth because the utility loss from foregoing consumption to extend life is lower when one is wealthier. Patent-protected monopolies may be able to extract this increased value of health and raise prices more and more for the same gains in health as incomes rise.

Another proposed explanation for increasing quality-adjusted prices in health care is that a large portion of care is paid for by third-party payers – either public or private – so demand is not sensitive to price. However, this is more of an explanation for why markups in health care may be high than for why they would be increasing (without an increase in health insurance coverage). Others have argued that third-party payers decrease the incentive for cost-reducing innovations (Weisbrod, 1991). However, patients ultimately have to demand and pay the resulting higher premiums. Moreover, third-party payers seem almost "hyper-rational" in their purchasing decisions in that they use the very same cost-effectiveness studies analyzed here. In very few other industries do buyers use explicit quantitative metrics like these to quantify the costs and quality of products before purchasing. Though reimbursement by payers based on cost-effectiveness is less institutionalized in the more privately financed US market than in Europe, the majority of cost-effectiveness studies are actually done for the US market and funded by US manufacturers that would be unlikely to fund them unless they influenced US payers.

In summary, we believe that more systematic inquiry is needed on the impact of technological change on quality-adjusted prices in health care. The central debate is around whether increased spending is justified by greater health benefits. We have argued that the cost-effectiveness literature has implicitly analyzed such quality-adjusted prices over the last half century, but the analysis of quality-adjusted prices by economists has not incorporated this large body of evidence.

References

Beasley, D. (2015). Gilead boosts hepatitis C drug discounts, shares slide. *Market News*. http://www.reuters.com/article/gilead-sciences-results-idUSL1N0VD2V920150203.

Cutler, D. M., McClellan, M., Newhouse, J. P., and Remler, D. (1998). Are medical prices declining? Evidence from heart attack treatments. *Quarterly Journal of Economics*, 113(4):991–1024.

Dow, W. H., Philipson, T. J., and Sala-i Martin, X. (1999). Longevity complementarities under competing risks. *American Economic Review*, 89(5):1358–1371.

Frank, R. G., Berndt, E. R., and Busch, S. H. (1999). Price indexes for the treatment of depression. In Triplett, J. E., editor, *Measuring the Prices of Medical Treatments*. Brookings Institution Press.

Hall, R. E. and Jones, C. I. (2007). The value of life and the rise in health spending. *The Quarterly Journal of Economics*, 122(1):39–72.

Hartung, D. M., Bourdette, D. N., Ahmed, S. M., and Whitham, R. H. (2015). The cost of multiple sclerosis drugs in the US and the pharmaceutical industry. *Neurology*, 84(21):2185–2192.

Howard, D. H., Bach, P. B., Berndt, E. R., and Conti, R. M. (2015). Pricing in the market for anticancer drugs. *The Journal of Economic Perspectives*, 29(1):139–162.

Jena, A. B. and Philipson, T. J. (2008). Cost-effectiveness analysis and innovation. *Journal of Health Economics*, 27(5):1224–1236.

Lucchinetti, C., Bruck, W., Parisi, J., Scheithauer, B., Rodriguez, M., Lassman, H., et al. (2000). Heterogeneity of multiple sclerosis lesions: Implications for the pathogenesis of demyelination. *Annals of neurology*, 47(6):707–717.

Newhouse, J. P. (1992). Medical care costs: How much welfare loss? *The Journal of Economic Perspectives*, 6(3):3–21.

Smith, B., Carson, S., Fu, R., McDonagh, M., Dana, T., Chan, B. K., Thakurta, S., and Gibler, A. (2010). Drug class review: Disease-modifying drugs for multiple sclerosis.

Weisbrod, B. A. (1991). The health care quadrilemma: An essay on technological change, insurance, quality of care, and cost containment. *Journal of Economic Literature*, 29(2):523–552.

Zhang, S., Bastian, N. D., and Griffin, P. M. (2015). Cost-effectiveness of sofosbuvir-based treatments for chronic hepatitis C in the US. *BMC Gastroenterology*, 15(1):1.