THE
QUARTERLY JOURNAL
OF ECONOMICS

Vol. CXXI February 2006 Issue 1

THE DISTORTIONARY EFFECTS OF GOVERNMENT PROCUREMENT: EVIDENCE FROM MEDICAID PRESCRIPTION DRUG PURCHASING*

MARK DUGGAN AND FIONA M. SCOTT MORTON

In 2003 the federal-state Medicaid program provided prescription drug coverage to more than 50 million people. To determine the price that it will pay for each drug, Medicaid uses the average private sector price. When Medicaid is a large part of the demand for a drug, this creates an incentive for its maker to increase prices for other health care consumers. Using drug utilization and expenditure data for the top 200 drugs in 1997 and in 2002, we investigate the relationship between the Medicaid market share (MMS) and the average price of a prescription. Our estimates imply that a 10-percentage-point increase in the MMS is associated with a 7 to 10 percent increase in the average price of a prescription. In addition, the Medicaid rules increase a firm's incentive to introduce new versions of a drug in order to raise price. We find empirical evidence that firms producing newer drugs with larger sales to Medicaid are more likely to introduce new versions. Taken together, our findings suggest that government procurement rules can alter equilibrium price and product proliferation in the private sector.

* The authors are grateful to the Merck Corporation for generously allowing us access to IMS data. We also thank IMS for permission to use their national sales data. The authors are grateful to the National Science Foundation for supporting this research. Duggan also thanks the Alfred P. Sloan and Robert Wood Johnson Foundations for support. We thank Sarah Bohn, Julian Cristia, Tamara Hayford, Jillian Popadak, and Sabrina Yusuf for excellent research assistance and are grateful to Alberto Abadie, Christopher Adams, Leemore Dafny, Edward Glaeser, Philip Haile, Ginger Jin, Andrei Shleifer, three anonymous referees, and seminar participants at Cornell University, the Federal Trade Commission, HEC Montreal, Massachusetts Institute of Technology, University of Michigan, Northwestern University, and Stanford University for helpful comments. The views expressed in the paper are those of the authors and not necessarily the views of any sponsoring organization. Contact information for the authors follows: Mark Duggan, Economics Department, University of Maryland, 3115L Tydings Hall, College Park, MD, 20742: *duggan@econ.umd.edu*. Fiona Scott Morton, Yale School of Management, 55 Hillhouse Avenue, New Haven, CT 06520: *fiona.scottmorton@yale.edu*. All errors are our own.

The Quarterly Journal of Economics, February 2006

I. Introduction

Does the Medicaid program distort the market price of pharmaceuticals and the number of versions of those drugs on the market? In 2003 the Medicaid program provided health insurance coverage for most health care services, including pharmaceuticals, to more than 50 million low-income people in the United States [Kaiser 2005]. There is a large body of research examining the effects of this program on labor supply, savings, health care utilization, insurance coverage, and health outcomes.[1] One issue that has received virtually no attention in previous work is the effect of Medicaid on the prices of health care treatments.[2]

The contribution of this paper is to demonstrate that the effects of Medicaid procurement of prescription drugs are important. The basic intuition is straightforward and can be conveyed with three facts. First, Medicaid currently accounts for approximately 19 percent of pharmaceutical spending in the United States, with this fraction varying substantially across drugs, and Medicaid covers essentially all drugs after their approval by the Food and Drug Administration (FDA). Second, the price Medicaid pays a manufacturer for its drug is a set fraction of the average price paid by non-Medicaid consumers. And finally, a large fraction of pharmaceutical products enjoy some degree of market power due to patent protection.

Thus, if a firm raises its price to non-Medicaid customers, it will receive a higher price for all Medicaid prescriptions filled. Because Medicaid recipients do not share in the cost of their medical care, they are likely to be unresponsive to changes in price. As government purchases become large, it is clear that linking prices in this way can create significant incentives to raise price in the private market.

To empirically test our hypothesis, we exploit variation in the importance of Medicaid purchasing across drugs. For example, in 2002 the program accounted for less than 6 percent of revenues for Lipitor, the top-selling drug in the United States. In that same year, nearly 75 percent of U. S. revenues for Zyprexa, the number

1. See papers such as Cutler and Gruber [1996] and Currie and Grogger [2002] on health care utilization, Epstein and Newhouse [1998] on health outcomes, Yelowitz [1995] on labor supply, and Gruber and Yelowitz [1999] on savings. For an excellent review of this literature, see Gruber [2003].

2. One exception is Scott Morton [1997] in which the author examines the effect of the most-favored-customer clause in Medicaid on price levels and price dispersion in the non-Medicaid market.

eight seller overall, came from the Medicaid program. Our key explanatory variable is the Medicaid market share, which is defined as Medicaid revenues for a drug in year t divided by total revenues in that same year. We explore the relationship of this variable with pharmaceutical prices while controlling for other observable factors that should influence this outcome variable of interest.

Focusing first on the top 200 drugs in 1997, our findings demonstrate that pharmaceutical prices are significantly positively related with the Medicaid market share. The point estimates in our baseline specifications imply that a 10-percentage-point increase in the Medicaid market share is associated with a 7 percent increase in pharmaceutical prices for the 1997 sample and a 10 percent increase for the 2002 sample.

One concern with this first set of estimates is that they may partially reflect an effect of prices on Medicaid market share rather than the reverse. Physicians may encourage Medicaid recipients to purchase expensive drugs while steering other patients with private insurance or no insurance to cheaper alternatives. To address this possible source of bias, we instrument for the Medicaid market share using variation across illnesses in the fraction of people enrolled in Medicaid, which we obtain from the National Ambulatory Medical Care Survey (NAMCS). Our results using this instrumental variables strategy are quite similar to those described above. We also estimate an analogous set of specifications for generic drugs, whose manufacturers should have much less power to set prices, and find no evidence of a positive relationship between Medicaid market share and price. These results strongly suggest that our estimates are capturing a causal effect of the Medicaid program on pharmaceutical prices.

In the second main part of our paper, we explore the effect of another potentially important feature of the Medicaid procurement rules: price increases for any particular treatment may be no greater than inflation as measured by the Consumer Price Index (CPI). If the optimal price for a drug increases more rapidly than this index, there is an incentive over time to introduce new versions of a drug with different dosage amounts or route types (e.g., capsule, liquid, tablet) that would have unrestricted base prices. Firms must weigh the resulting increase in revenue against the cost of filing a new drug application with the FDA. Theoretically, one would expect firms with relatively large current and expected future sales to Medicaid to be most affected by this regulation. Consistent with this, our results demonstrate that new drugs with high Medicaid revenues introduce more new versions than do otherwise similar drugs.

Taken together, our results strongly suggest that Medicaid coverage of prescription drugs has increased the price paid by other health care consumers for these same treatments. Using our point estimates along with Medicaid shares for each of the top 200 treatments, the average price of a non-Medicaid prescription would have been 13.3 percent lower in 2002 in the absence of Medicaid's pricing rule.

The outline of the paper is as follows. Section II provides background on the Medicaid rules for determining pharmaceutical prices, formally lays out a pharmaceutical firm's financial incentives, and describes the data used in our empirical analyses. Section III presents results for the effect of the Medicaid market share on average pharmaceutical prices, while Section IV describes our results for the effect on the rate at which new versions of a drug are introduced. The final section concludes and discusses important directions for future work.

II. Institutions, Incentives, and Data

II.A. Medicaid's Rules for Setting Pharmaceutical Prices

Medicaid provides insurance for all major categories of medical care including prescription drugs. In 2003, this means-tested program provided coverage to more than 50 million people and accounted for $280 billion in federal and state government spending. In contrast to Medicare, the program has almost no co-pays[3] or deductibles, and thus beneficiaries have little incentive to consider cost when choosing between treatments [Duggan 2005].

Each state administers its own Medicaid program but uses essentially the same formula for setting prices for each one of the thousands of covered drug treatments. The price for a brand name (as opposed to a generic) prescription drug is determined as follows. First, the state government reimburses the pharmacy that dispensed the drug according to a formula determined at the state level. The usual approach to reimbursement is some fraction of the drug's AWP, which is a standard industry list price. This scaling factor varies from a low of 86.5 percent in Michigan to a high of 95 percent in Alaska [Kaiser 2001]. The pharmacy also receives a small dispensing fee that is approximately $5.00 in

3. In recent years a number of states have introduced most co-pays. The largest co-pay in any state in the first quarter of 2005 is $3 and in most states this does not vary across brand drugs. Only seven states have a co-pay that varies with price, with the maximum range from $0.50 to $3.00.

the typical state. Next, the federal Centers for Medicare and Medicaid Services (CMS) calculates a rebate that is equal either to 15.1 percent of the average manufacturer price (AMP) or to the difference between the AMP and the lowest price given to any buyer in the United States, whichever is greater.[4] AMP is the average price at which the manufacturer sold the drug, and this is reported to CMS.[5] This rebate amount is multiplied by the quantity purchased by each state and paid by the manufacturer to each state program so that the effective cost of drugs to the state is reduced.

Additionally, the program specifies that after a new drug treatment has entered the market, its Medicaid price cannot increase more rapidly than the rate of inflation as measured by the Consumer Price Index (CPI). If a firm's optimal price for a drug is increasing more rapidly than the CPI, it can raise its price, but Medicaid will not pay the portion of the price increase that is greater than the CPI.[6] However, the CPI cap applies to changes in price of the *same product* over time. The level at which a product is defined for purposes of Medicaid rebates is the first nine digits of the national drug code (NDC). An NDC is an eleven-digit code that defines a product perfectly: drug, form, route, strength, and package. The last two digits determine package size and are not used by the Medicaid rules (so a new package does not count as a new product). If a manufacturer alters strength, form, or route, it will be assigned a new NDC for the product, which may be priced in any way the firm desires. The rate of increase from that base is what is compared with CPI growth going forward.

Thus, a firm that wants to increase its Medicaid prices faster than the CPI can avoid the constraint by introducing new versions of its drug. This kind of behavior was first documented in

4. Thus, the net Medicaid price for a drug with AWP = $110 and AMP = $90 in a state that paid 90 percent of AWP to its pharmacies would be less than or equal to $85.41 (=(.90 ∗ 110) − (.151 ∗ 90)). If a firm sold to a non-Medicaid customer in the United States at a price less than $85.41 during that quarter, then the net Medicaid price would fall below this amount. There are a number of exceptions to this general description. For example, very low prices given to teaching hospitals, clinics, or other government programs do not count as the "best price."

5. The definition of AMP used by CMS is "the average price at which a manufacturer sells a particular drug to purchasers, not including sales to federal purchasers or state drug assistance programs" [Aldridge and Doyle 2002].

6. The way this works is as follows. First CMS determines whether the best price or the straight 15.1 percent discount applies and determines the Medicaid rebate. Then it calculates whether the increase in its AMP since launch is greater than the increase in the CPI during that time. If so, the firm's rebate increases by the difference between the actual and maximum allowable AMP. See *http://www.cms.hhs.gov/medicaid/drugs/drug12.asp* for more details.

the regulation literature by Averch and Johnson [1962] and later by many others (e.g., Parker [1999], Borrell [1999], Olson [1996], and Laffont and Tirole [1993]). The distortion of activity by the regulated entity is akin to the theoretical effects found in the principal-agent literature on multitasking [Holmstrom and Milgrom 1991], in which the agent's (firm's) measured action responds to incentives, but this is counteracted by an unmeasured action. In these papers the firm responds to the constraints or payoffs for a specific task, which in our case is keeping price increases on existing medicines below a threshold. It uses another activity (introducing new versions of drugs) that is not specified in the regulation to mitigate the cost of the first constraint. This literature demonstrates the difficulty in achieving the regulator's goals when those goals are in conflict with firm profitability and the manager has a choice of activities. For example, research on the Japanese market for prescription drugs documents frequent product introductions at high prices; and ascribes this behavior to regulation that allows considerable freedom to the manufacturer in setting the initial price but then imposes steep discounts after launch [Ikegami, Ikeda, and Kawai 1998].

II.B. Procurement Incentives

Consider two drugs that face identical demand curves except that part of the consumer base of one drug is eligible for the Medicaid program. Also assume that once enrolled in the program, Medicaid recipients have perfectly inelastic demand for the drug.[7] Finally, the firm is not permitted to sell only to the Medicaid segment, but must sell into the private market and set a uniform price across segments. In this situation the firm with Medicaid consumers will charge a higher price than the other, due to its more inelastic demand curve. Thus, by tying the Medicaid price to the average price in the private sector, this procurement rule adversely affects other health care consumers through its effect on price. This result is not specific to Medicaid but holds in the more general case of a monopolist that is constrained to

7. Since Medicaid recipients typically do not share in the cost of their prescription drugs, it is reasonable to assume that the actual price the firm charges the government does not affect their demand. However, many state governments have policies that attempt to steer Medicaid recipients toward lower cost prescriptions. These focus almost exclusively on physicians and other health care providers by, for example, requiring them to prescribe a generic version when it is available. However, providers can usually get around these constraints by, for example, writing "brand medically necessary" on the prescription.

choose the same price in market segments with different elasticities [Tirole 1991, pp. 137–138].

As mentioned above, the Medicaid program constrains the rate at which the price of a particular treatment can increase over time to be no greater than inflation as measured by the CPI. If the profit-maximizing price for a drug is increasing more rapidly than this index, then a firm could respond by introducing new versions of the same drug and price the new versions higher than the old versions. Of course, in order to place these new products on the market, the firm must receive approval from the FDA. Approval for a new version of an already approved drug requires either a supplemental new drug application (NDA) or a completely new NDA, depending on the significance of the change.[8] The costs of getting approval for a new version of an existing drug are the fee paid to the FDA, the cost of the clinical trials, and the delay.[9] The firm would then weigh these costs against the benefits of unconstrained pricing.

The benefit of unconstrained pricing relative to constrained pricing is that the firm can increase its price to Medicaid when it introduces a new version. This is particularly important if the firm's initial price becomes suboptimal (too low) due to unexpected demand growth from either the private or public sectors. If, on the other hand, the firm can perfectly forecast demand growth in excess of the CPI, it still has an alternative margin on which to respond. The firm can price high initially in order to relax the Medicaid constraint in the future. This strategy is also clearly costly, as it involves suboptimal prices being charged for an extended period of time, and the risk that the environment changes and the chosen fixed price is not even the constrained optimum.

8. Changes that are expected to have minimal chemical effects can be approved with a supplemental NDA, which does not require any clinical trials and so is relatively inexpensive. Minor changes include a different strength of an already approved medication or a change in manufacturing location. Any change in dosage form or in the recommended dosage amount will require a new NDA supported by Phase III clinical trials that show the new form of the drug achieves the same safety and efficacy standards as the original form.

9. Average delays for a new drug application (NDA) are 10 months and for supplemental NDAs 4–6 months. Fees required by the FDA for either type of application averaged approximately $500,000 during our study period. The approximate cost of trials can be measured in the cost per patient recruited into a study, and the number of patients needed to show that the effects alleged by the applicant are statistically significant. Informal estimates of the costs of clinical trials range from $500 to $10,000 per patient depending on the length of time required to demonstrate results. For example, a cardiovascular drug would have to be taken longer (several months) than an antibiotic (several days) to demonstrate efficacy.

In an environment of rising prices, the procurement rules create an incentive such that when sales to Medicaid are greater, the firm will find it more attractive to introduce a new version of its drug. Thus, one would expect drugs with high sales to Medicaid—all else equal—to introduce more new versions in response to the program's CPI constraint.

II.C. Medicaid and IMS Utilization Data

We merge together data from two main sources. The first source was provided by IMS Health and contains total sales of all prescription drugs nationally from 1997–2002 for each combination of manufacturer and product. The data divide drug treatments into fourteen therapeutic classes such as mental disorders, cardiovascular illnesses, and the alimentary tract. Within each of these therapeutic classes, drugs are further divided into 249 subclasses, such as cholesterol reducers, antiulcerants, and antidepressants.

There are 2391 different ingredient names in the data with strictly positive sales in one or more years between 1997 and 2002. An ingredient defines just one drug, though for many drugs there are multiple products in the IMS data. For example, a generic introduced after the patent for a brand name drug expires would be a different product. After the patent on Prozac expired in 2001, nearly twenty firms began to produce its generic equivalent fluoxetine, each of which is a separate product in our data. Additionally, a firm may release a different product with patent protection, such as an extended release version, for its own branded drug.[10] In the same year that Prozac's patent expired, it released a new version called Prozac Weekly that patients took just once per week. This appears as a separate product in our IMS data, but both the extended release version of Prozac and its generic versions all have the same ingredient name in our data.

An observation is a drug and not a product because a product is the outcome of proliferation. Our measure of drug revenue therefore adds together the revenue of different products within a drug, all made by the innovator. This revenue measure does not include sales of generic competitors produced by different firms. Thus, in the example above, sales for both Prozac and Prozac Weekly would be included in U. S. revenues for Prozac, but those for the generic versions of fluoxetine would not.

10. In the IMS data, changes in dosage amount or route type usually do not show up as a separate product.

Our second main source of data is the CMS Drug Utilization files which track Medicaid spending by product code (NDC) in all 50 states and the District of Columbia in each year. In 1997 there were 26,130 nine-digit NDCs that had nonzero Medicaid utilization with total spending equal to $13.5 billion.[11] Thus, Medicaid accounted for 16.9 percent of the $79.9 billion in prescription drug spending in the United States during that year. In 2002 the number of NDCs was 22,111, and total Medicaid prescription drug spending was $29.6 billion, or 18.5 percent of the $160.6 in drug spending in the United States.[12] The Medicaid data also include variables that allow us to determine whether a particular version of a drug is available over-the-counter or if it is generic, along with the FDA approval date and the date that each version entered the market. We define the number of NDCs for each drug in year t as equaling the number with more than 0.1 percent of each drug's total prescriptions in that year. We define a version as entering between 1997 and 2002 if it had zero spending in 1997 and had more than 0.1 percent of the drug's sales in 2002.[13]

We then match the Medicaid and IMS National Sales Perspective data using drug names.[14] Having done this, we can calculate a Medicaid market share by dividing Medicaid expenditure in all states by total U. S. sales. In our empirical analyses

11. Dollar figures cited in the paper are adjusted to 2002 dollars using a 1997 CPI of 160.5 and a 2002 CPI of 179.9.

12. Note that our Medicaid expenditures do not include rebates paid to state governments, which accounted for approximately 15 percent of gross spending on prescription drugs in 1997 and in 2002 and which should be subtracted from these expenditure figures. Despite repeated requests to CMS, we were unable to obtain these data because they consider them proprietary. The fact that this 15 percent number is very close to the 15.1 percent of AMP described above suggests that there is not too much heterogeneity across drugs in the rebate amount as a fraction of the gross price. The absence of rebate data suggests that we will to some extent overstate the Medicaid market share. On the other hand, our data do not include spending under the 340B program of the Veterans Healthcare Act of 1992 (VHCA) which effectively allows some public health service programs to join the Medicaid procurement program. For example, state ADAP programs (AIDS drug assistance, $250 million in spending in 1997) may get their drugs this way (see Aldridge and Doyle [2002]). We do not have purchase data from these organizations, but as long as they are not both large *and* disproportionately purchasing non-Medicaid drugs, then we simply have a countervailing understatement of our Medicaid expenditure measure.

13. See Table 1 in Duggan and Scott Morton [2004] for utilization and sales data for the ten different versions of Zyprexa, the drug with more Medicaid spending than any other in the 2002 calendar year.

14. The names are rarely identical between the two files because the Medicaid names often provide details on dosage amount or route type. Thus, we take care to capture all Medicaid NDCs that contain the string of each of the 200 names in our IMS data. This involved a substantial amount of hand-checking because there were some spelling errors in the Medicaid data (e.g., Prilopec instead of Prilosec). Of the thousands of NDCs in each year, fewer than 5 percent are in the top 200 drugs sold in the United States.

we focus on the top 200 drugs in 1997 and in 2002 in terms of total U. S. revenues in each year. Our primary reason for limiting our focus is that we also use data from a publication called *Med Ad News.* This magazine publishes an annual list of the revenues of the top 200 global drugs and lists which conditions each drug is used for. These conditions form part of the input into our empirical analyses below. Additionally, the drugs in the top 200 include most of the spending on pharmaceuticals in the United States: our 1997 sample of 200 drugs accounts for 64.1 percent of total U. S. sales in IMS National Sales Perspectives (TM) data, while those in our 2002 sample are responsible for 72.6 percent of the total 2002 spending. Because the IMS National Sales Perspectives (TM) data also include sales of generic and over-the-counter drugs, the share of spending on brand-name drugs accounted for by the 200 drugs in our 1997 and 2002 samples are closer to 75 and 80 percent, respectively. So we are capturing the majority of the market with these 200 top-selling drugs.

In our examination of prices and proliferation, we are interested in the incentives the regulation creates for brand name prescription drugs. We therefore exclude over-the-counter drugs because these are paid for out of pocket and thus are not typically covered by Medicaid. We also exclude generic drugs because they face different Medicaid regulations and because the producers of these drugs will have less pricing power (because of free entry and homogeneity). We do not ignore generic drugs, however, as we include measures of whether our brand-name drugs face generic competition in our empirical analyses of pricing and proliferation below. To account for the amount of competition that each drug faces, we also consider the number of therapeutic substitutes for each drug. To estimate this, we calculate the number of other molecules within each drug's subclass with nonzero sales in the relevant year. More details on the construction of these and other variables are included in the Appendix.[15]

Table I provides summary statistics for our top 200 drugs in 1997 and 2002. The average drug in our 1997 sample has U. S. revenues of $256 million and a Medicaid market share of 16.5 percent. The corresponding figures for the 2002 sample are $583

15. In some cases a drug falls into more than one therapeutic category or more than one subclass. In these cases we assign the drug to the category or subclass with the highest expenditures. This is possible because the IMS data differentiate between uses of a drug, and thus, one can determine what fraction of revenues for the drug fall into each category or subclass.

TABLE I
SUMMARY STATISTICS FOR 1997 AND 2002 SAMPLES OF TOP 200 DRUGS

1997 Sample	Mean	Std. dev.	Median	Min.	Max.
Total revenues (/1000)	255756	307930	138250	53708	2188413
Medicaid revenues (/1000)	43274	65161	20632	3	413699
Medicaid market share	0.165	0.151	0.136	0.000	1.000
Predicted Medicaid share	0.085	0.063	0.070	0.000	0.394
Price per prescription	141	223	63	18	1645
Log (price per prescription)	4.39	0.90	4.14	2.87	7.41
Price rank	100.5	57.9	100.5	1	200
# of substitutes	14.13	13.58	11	0	146
Any generic competition	0.280	0.450	0	0	1
Years on the market	9.10	8.30	6	0	55
# of NDCs in 1997	3.73	2.52	3	1	15
# of NDCs in 2002	4.12	2.63	4	1	17
# of new NDCs	0.58	1.03	0	0	6
2002 Sample	Mean	Std. dev.	Median	Min.	Max.
Total revenues (/1000)	583024	718636	31520	117039	5677668
Medicaid revenues (/1000)	102186	172950	50029	108	1668500
Medicaid market share	0.177	0.150	0.140	0.000	1.000
Predicted Medicaid share	0.092	0.075	0.071	0.000	0.394
Price per prescription	280	449	93	20	2883
Log (price per prescription)	4.90	1.09	4.53	3.01	7.97
Price rank	100.5	100.5	100.5	1	200
# of substitutes	14.99	12.17	13	0	127
Any generic competition	0.165	0.372	0	0	1
Years on the market	8.23	7.48	6	0	60

Table I provides summary statistics for the top 200 drugs in 1997 and in 2002 in terms of total revenues in the United States. Generic and over-the-counter drugs are excluded. Expenditure amounts are adjusted to 2002 dollars using the Consumer Price Index for all urban consumers (CPI-U).

million and 17.7 percent.[16] The variable *Price per Prescription* is equal to total Medicaid revenues divided by Medicaid prescriptions in each year. We use the Medicaid data to calculate price because it is not available in the IMS data (though a comparison of this with data from another source suggests that it is an accurate measure of average private prices in the United States).[17] To reduce undesirable skewness in this measure, we

16. For one drug in 1997 and three in 2002 Medicaid revenues exceed reported U. S. revenues, and we therefore set the Medicaid share equal to 1.00. Our results were not substantially affected if we exclude these drugs from our subsequent empirical analyses.

17. We would prefer to use the average price of a private-sector U. S. prescription as our measure of price, rather than a measure based on Medicaid purchases, but this information is unfortunately not publicly available and is not included in our IMS data. To investigate whether our measure of price is accurate, we compared our data

take the log of the average price and report this with the other summary statistics.

Table I also summarizes measures of competition for each of the top 200 drugs. The variable # *of Substitutes* is equal to the number of ingredients (excluding the drug itself) within the drug's IMS subclass in the relevant year. These competitors are therapeutic substitutes for the focal drug. For both samples the average of this variable is approximately fourteen. There is a substantial difference between the two samples in the fraction of drugs facing generic competition, which declines from 28 percent in the 1997 sample to 16.5 percent in the 2002 sample. This is perhaps partly because the average number of years that each drug is on the market falls from 9.1 to 8.2 and thus a smaller fraction of patents would have expired in the more recent sample.

III. The Effect of Medicaid Market Share on Pharmaceutical Prices

In this section we investigate the effect of a drug's Medicaid market share on its average price. To estimate the importance of this effect, one would ideally use sharp and unanticipated changes in the Medicaid market share for one or more prescription drugs. This would allow one to control for time-invariant factors that are specific to each drug, including its effectiveness, ease of use, and production costs. Unfortunately, no such changes exist because Medicaid eligibility does not change rapidly from one period to the next, and because Medicaid covers a drug once it receives FDA approval.[18] Later in this section we utilize an

with Scott-Levin data reported for the top 50 drugs. (Scott-Levin is the other main private sector source of pharmaceutical expenditure and utilization data.) A regression of the log of the average price of a Medicaid prescription on the log of the Scott-Levin measure yields a coefficient of 1.036 with a standard error of just .024. It therefore appears that our Medicaid data provide a good estimate of average pharmaceutical prices for all health care consumers in the United States. Our data are from 2002; the Scott-Levin data are from a 2001 and appeared in a publication by the National Institute of Health Care Management [2002]. Five of their fifty drugs are not included in the regression because they are generic. Recall that rebates are excluded from our measure of Medicaid revenues.

18. Plausibly exogenous increases in Medicaid share are not easy to come by. Aged, blind, and disabled recipients of Supplemental Security Income benefits account for the vast majority of pharmaceutical spending. For example, in California in 1997 more than 81 percent of prescription drug spending was for beneficiaries of this program (despite being less than one-fourth of recipients). The number receiving SSI nationally was fairly steady, increasing by less than 5 percent from 6.495 million to 6.788 million during our 1997–2002 sample period. There were big increases from the mid-1980s to the mid-1990s, but we do not have data for this earlier period. Increases in Medicaid coverage for children will not

instrumental variables strategy to account for the endogeneity of the Medicaid market share.

We begin by exploiting the considerable variation across drugs in the Medicaid market share (*MMS*) to estimate the effect of the program on pharmaceutical prices. As our measure of *MMS* we take the ratio of Medicaid revenues to total U. S. revenues for that same drug. We control for other factors that are likely to influence the profit-maximizing price for a drug. The existence of generics introduces substantially more, and different, competition into the market, so we include an indicator of whether or not there is generic competition (*GenComp*) in the specification.[19] As more therapeutic substitutes (*Subst*) become available, the drug will face additional competition and its optimal price should decline. Additionally, newer drugs tend to be more expensive than earlier ones, because they embody more new technology, and thus it is important to control for the time (*Years*) that a drug has been on the market. And finally, we include indicator variables for each of the fourteen major therapeutic categories when estimating specifications of the following type:[20]

$$(1)\quad \log(Price_{jt}) = \alpha + \beta_1 MMS_{jt} + \beta_2 GenComp_{jt} + \beta_3 Subst_{jt} + \beta_4 Years_{jt} + \sum_{k=1}^{14} \theta_k^* I(Class_{jt} = k).$$

give us much increase in Medicaid market share both because they do not consume many drugs and because they will often end up in Medicaid-managed-care plans (if not on SSI). These plans typically buy the drugs themselves and thus do not contribute to our measured Medicaid market share (unless the plans use Medicaid prices, in which case they would appear in the state drug utilization data).

19. The literature is divided on whether generic competition a) lowers the optimal price for a brand-name drug as consumers can easily substitute to a cheaper version, or b) raises the price of the brand as all elastic demanders move to the generic and the remaining customers have inelastic demand [Frank and Salkever 1997, 1992]. Our test of interest (MMS) is not affected by which strategy brands use.

20. The fourteen category fixed effects control for broad differences across drugs in, for example, the share of privately insured consumers and the patients' demographic characteristics. The classes are quite general. For example, there is one category for drugs used to treat mental disorders, another for drugs used to treat cardiovascular illnesses, and another for drugs used to treat cancer. There is considerable variation in the Medicaid market share within the fourteen therapeutic classes. A regression of the Medicaid share on the fourteen indicator dummies produces an R^2 of just 0.22 in 1997 and 0.30 in 2002. The weighted average standard deviation within a therapeutic category (weighting each category's standard deviations by its share of the top 200 drugs) is about .11 in both years. This compares with overall standard deviations of .15 in both years. Thus, there is somewhat less variation within therapeutic categories than across all drugs, but not much less.

In this equation, the dependent variable is equal to the log of the average price for a prescription for drug j in year t. Because even a log transformation of the average price is skewed [Duggan and Scott Morton 2004], we use rank order as an alternative dependent variable.

The coefficient of particular interest in equation (1) is β_1, which captures the conditional relationship between the Medicaid market share and price. In Table II we summarize the results of specifications of equation (1) for data from 1997 and 2002. Our results for the top 200 drugs in 1997 and 2002 are in the odd-numbered columns of the table. The point estimates of 0.674 and 0.943 in columns (1) and (3) are significant at the 5 percent level and imply that a 10-percentage point increase in the Medicaid share is associated with approximately a 7.0 percent increase in the average price in 1997 and a 10.0 percent increase in 2002. Drugs purchased disproportionately by Medicaid are significantly more expensive than otherwise similar drugs. The estimated coefficient of *Years* is negative and significant as expected. The estimate for the presence of generic competition is negative in both years but significant only in 1997. This negative relationship may actually be capturing an effect of the age of the drug, as the number of years on the market does not measure age exactly. The number of therapeutic substitutes has a significantly negative estimate in the specifications for both 1997 and 2002.

Because of skewness in the distribution of the log of pharmaceutical prices, we next estimate specifications in which we use each drug's rank in the price distribution as the dependent variable. This variable is uniformly distributed from 1 to 200 with the most expensive drug having a value of 200. The results using this alternative measure are summarized in columns (5) and (7) of Table II and are qualitatively similar to the originals. In both cases the estimates are statistically significant, at the 5 percent level in 1997 and at the 1 percent level in 2002. Notice that in both specifications the implied effect of the Medicaid market share is substantially larger in 2002, suggesting that the effect of Medicaid may be increasing over time.

We now turn to robustness checks. One concern with this first set of estimates is that it may be capturing an effect of price on Medicaid rather than the reverse. For example, a physician may steer Medicaid recipients to an expensive drug while encouraging those with private insurance or no insurance to purchase a cheaper alternative within the same therapeutic class. If this reverse causation were important, it would bias up our estimates

TABLE II
OLS ESTIMATES OF THE IMPACT OF MEDICAID MARKET SHARE ON PHARMACEUTICAL PRICES

	Log(price)				Price rank			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Medicaid market share	0.674**	1.827**	0.943**	2.044***	41.6**	44.1**	65.4***	39.5**
	(.317)	(.837)	(.410)	(.750)	(17.7)	(20.5)	(20.5)	(15.7)
Any generic comp	−0.231**	−0.289	−0.095	−0.403	−13.6**	−10.7	−9.3	−10.6
	(.088)	(.193)	(.144)	(.256)	(6.7)	(6.6)	(8.6)	(6.8)
# of substitutes	−0.008**	−0.004	−.011*	−.007*	−0.578**	−0.176	−.767*	−.260**
	(.003)	(.004)	(.007)	(.004)	(.239)	(.165)	(.412)	(.112)
Years on market	−0.020***	−.023***	−0.028***	−.035***	−1.23***	−.693**	−1.48***	−.938***
	(.006)	(.008)	(.009)	(.011)	(.44)	(.279)	(.47)	(.264)
# observations	200	126	200	125	200	126	200	125
Sample	1997	1997	2002	2002	1997	1997	2002	2002
Class fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.593	0.389	0.621	0.515	0.556	0.381	0.567	0.517

Specifications in odd-numbered columns include the top 200 drugs in terms of total revenues in either 1997 or 2002. Specifications in even-numbered columns include the top-selling drug in each therapeutic subcategory in 1997 or 2002. All four samples exclude over-the-counter and generic drugs. The dependent variable in the first four columns is the log of the average price per prescription, and in the second four columns is the price rank (200 is the most expensive drug and 1 is least expensive). The Medicaid market share for each drug is equal to Medicaid revenues divided by total revenues in the United States. Each specification includes fourteen therapeutic class fixed effects. Huber-White standard errors are reported in parentheses. ***, **, and * represent significance at the 1, 5, and 10 percent levels, respectively.

for the effect of MMS. To address this possibility, we estimated a companion set of specifications that considered only one branded drug from each of the 242 subcategories with nonzero sales in each year. The resulting variation in Medicaid market share across these drugs should largely be driven by differences in the fraction of individuals with particular illnesses on Medicaid. In some categories the top selling drug is either generic or available over-the-counter, so our sample has approximately 125 drugs in each year. The results for this second set of specifications are summarized in the even-numbered columns of Table II and present a quite similar picture. In all four specifications the Medicaid market share is significantly positively related with average pharmaceutical prices.[21]

Another possible source of bias is the presence of an omitted variable that is correlated with the Medicaid market share and also influences firms' profit-maximizing prices. For example, drugs consumed differentially by Medicaid recipients may have higher production costs than otherwise similar drugs. To examine this possibility, we estimated an analogous set of specifications on a sample of the 200 top-selling generic drugs in 1997 and 2002. Because producers of generic drugs would have much less ability to set prices, this set of specifications serves as a falsification test for the ones presented above. The results of these analyses are summarized in Table III. In no case is there a significant relationship between Medicaid market share and pharmaceutical prices. In fact, the estimates are negative in all four specifications, suggesting that if anything omitted factors would bias against finding a positive relationship between Medicaid market share and average prices for brand name drugs.

One potentially important issue is the sensitivity of our results to the exclusion of drugs with particularly high Medicaid market shares. As recent research has shown [Duggan 2005], drugs used to treat antipsychotic illness have the highest Medicaid market shares, with this program accounting for almost 75 percent of revenues for Zyprexa, Risperdal, Seroquel, Clozaril, and Geodon. Given that these few drugs have both a large MMS and are substantially more expensive than the average treatment, it is possible that the inclusion in our sample is to some extent driving our results. We investigate the sensitivity of our

21. We also interacted our measure of therapeutic substitutes with Medicaid market share to see whether drugs that face more competition respond differentially to the pricing rules. We use the same specification as the odd columns of Table II, and find that the coefficient is statistically insignificant.

TABLE III
OLS ESTIMATES OF THE IMPACT OF MEDICAID MARKET SHARE FOR GENERIC DRUGS

	Log(price)		Price rank	
	(1)	(2)	(3)	(4)
Medicaid market share	−0.081	−0.249	−17.8	−20.3
	(.263)	(.268)	(14.4)	(14.0)
# of substitutes	0.002	−.008*	0.272	−0.420
	(.004)	(.005)	(.249)	(.304)
Years on market	−.026***	−.040***	−1.68***	−2.37***
	(.008)	(.006)	(.48)	(.34)
# of firms producing drug	−.057***	−.054***	−3.32***	−3.18***
	(.009)	(.012)	(.54)	(.71)
# Observations	200	200	200	200
Sample	1997	2002	1997	2002
Class fixed effects?	Yes	Yes	Yes	Yes
R^2	0.519	0.507	0.469	0.453

Specifications include the top 200 generic drugs in terms of total revenues in either 1997 or 2002. The dependent variable in the first two columns is the log of the average price per prescription and in the second two columns is the price rank (200 is the most expensive drug and 1 is least expensive). The Medicaid market share for each drug is equal to Medicaid revenues divided by total revenues in the United States. Each specification includes fourteen therapeutic class fixed effects. Huber-White standard errors are reported in parentheses. ***, **, and * represent significance at the 1, 5, and 10 percent levels, respectively.

estimates to excluding all antipsychotic drugs in each year (3 in 1997, 4 in 2002) and find that our estimate for the effect of the Medicaid market share is essentially unchanged.

Another category of drugs with very high Medicaid market shares is HIV antiretroviral treatments. Almost half of U. S. residents with HIV are enrolled in the Medicaid program [Duggan and Evans 2005]. Our 1997 sample includes five drugs used to treat HIV, while our 2002 sample includes nine. Excluding any one of these drugs from our specifications has virtually no impact on our estimates but excluding all of them does have an effect. For example, in 2002 our baseline estimates of .943 and 65.4 for the log price and price rank specifications decline to .637 and 50.1, respectively, when all nine of the HIV/AIDS drugs are excluded.

Our results continue to hold if we exclude the HIV and antipsychotic drugs from our smaller sample of one drug per subcategory. For example, the baseline estimate of 39.5 in column (8) of Table II declines to 38.6 when we exclude Combivir and Zyprexa, the top-selling HIV and antipsychotic treatments, respectively, in 2002. A similarly small change occurs for our 1997 estimates when we exclude Epivir and Risperdal (the top HIV and antipsychotic drugs in that year). Taken together, these robustness checks therefore sug-

gest that our results are not solely attributable to the categories of drugs with the very highest Medicaid market shares.

As a final check on our estimates, we employ an instrumental variables strategy. We create a predicted MMS for each drug by estimating the fraction of potential customers of a drug who are insured by Medicaid rather than the actual fraction that purchases the drug (which is the Medicaid market share described above). To do this, we first use the list of the top 200 drugs in each year reported in the publication *Med Ad News* to determine for which conditions each drug is used. We then map this into diagnosis codes (ICD-9) using a comprehensive listing provided on the Centers for Disease Control (CDC) web site.[22] The third step is to determine the fraction of individuals with each diagnosis who are insured by Medicaid. To estimate this, we use the CDC's National Ambulatory Medical Care Survey (NAMCS) for 1997 and 2002, which provides detailed information about a large sample of patients in each year. We use the data on diagnosis and primary expected payer for 53,453 patients to count up how many have each diagnosis code as their primary diagnosis, and how many of them who have Medicaid as their primary expected payer. Aggregating these two numbers across diagnoses for each drug, we calculate our predicted Medicaid market share as the fraction of patients with the relevant diagnoses who are insured by Medicaid.[23] Note that if multiple drugs treat the same diagnosis, they will have the same predicted Medicaid share.

Columns (1) and (4) of Table IV summarize the results of our first-stage specifications. As can be seen in both columns, we find a statistically significant effect in the first stage: the predicted Medicaid share is significantly related to actual Medicaid share with a t-statistic greater than 10 in both 1997 an 2002. The remaining columns of Table IV display the results of the second stage with clustering by diagnosis, given that predicted MMS does not vary within diagnosis. We estimate similar or larger positive point estimates than OLS for the effect of Medicaid on prices in both the log price and the rank order specifications. However, the MMS coefficients in the 1997 sample are not sta-

22. See *ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/ICD-9/ucod.txt*

23. This predicted share will to some extent underestimate the Medicaid share because individuals who are dually eligible for both Medicare and Medicaid will be likely to report Medicare as their primary expected payer. While "dual eligibles" account for just 15 percent of Medicaid recipients, they account for almost 60 percent of Medicaid prescription drug spending. The mismeasurement of drug coverage for dual eligibles may partially explain why our first-stage estimates in Table IV are substantially greater than 1.0.

TABLE IV
ESTIMATES OF THE EFFECT OF MEDICAID MARKET SHARE ON PHARMACEUTICAL PRICES

	1997 Sample			2002 Sample		
	(1) MMS	(2) Log(price)	(3) Price rank	(4) MMS	(5) Log(price)	(6) Price rank
Medicaid market share		1.134	24.4		1.793*	92.3*
		(.853)	(44.4)		(.928)	(46.4)
Predicted MMS	1.697***			1.391***		
	(.158)			(.110)		
Any generic competition	−0.005	−0.189	−11.3	0.005	−0.111	−9.5
	(.019)	(.130)	(10.3)	(.019)	(.161)	(9.4)
# of substitutes	0.0003	−.008**	−.586**	0.0004	−0.003	−0.407
	(.0007)	(.003)	(.241)	(.0009)	(.008)	(.504)
Years on market	−0.0002	−0.021***	−1.33***	0.0012	−.023**	−1.21**
	(.0010)	(.008)	(.48)	(.0010)	(.009)	(.50)
# observations	187	187	187	177	177	177
Sample	1997	1997	1997	2002	2002	2002
Class fixed effects?	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.553	–	–	0.660	–	–

Specifications in each column include the 200 top-selling drugs in 1997 or 2002, though those for which predicted MMS is missing are excluded. The results from the first-stage specifications for the Medicaid market share (MMS) in 1997 and 2002 are summarized in columns (1) and (4), respectively. Predicted MMS is equal to the fraction of potential customers for a drug insured by Medicaid. Columns (2), (3), (5), and (6) summarize the results from specifications in which MMS is instrumented by the predicted Medicaid market share. The dependent variable in specifications (2) and (5) is the log of the average price per prescription and in specifications (3) and (6) is the price rank (200 is the most expensive drug and 1 is least expensive). MMS is the fraction of a drug's U. S. revenues accounted for by Medicaid. Each specification includes fourteen therapeutic class fixed effects. Standard errors are clustered by drug indication and are reported in parentheses. ***, **, and * represent significance at the 1, 5, and 10 percent levels, respectively.

tistically significant. We conclude that an instrumental variables approach yields similar results to the OLS specification.[24]

This first set of results strongly suggests that the rules used by the federal government to set prices for prescription drugs have led to an increase in prices faced by other health care consumers. This is consistent with the theoretical predictions, which suggest that the elastic segment of the market would be made worse off by this type of price-setting rule [Tirole 1991].

24. Our specifications assume that variation in drug prices does not cause people to become eligible for Medicaid. If individuals using expensive drugs qualify for Medicaid because of the high drug prices, then the estimates for β_1 would be biased as a result of reverse causation. Given that less than 10 percent of Medicaid recipients qualify for the program because of high medical expenses (through the Medically Needy and Medically Indigent programs) and given that prescription drugs account for just 12 percent of all Medicaid spending, this seems unlikely to be a significant factor. Additionally, SSI recipients—who would not qualify because of medical expenses but instead because of a disability—account for the vast majority of Medicaid spending on prescription drugs.

IV. The Effect of the Medicaid Market Share on New Drug Introductions

In this section we investigate whether drugs that sell disproportionately to the Medicaid program had more new drug introductions than their counterparts after controlling for other likely determinants of this outcome variable.

IV.A. Are New Versions of Prescription Drugs More Expensive?

In the Appendix we describe the results from specifications that investigate whether new versions of a drug are indeed more expensive than earlier ones. If they were not, then our hypothesis that Medicaid's CPI constraint creates an important distortion would be less plausible. The findings presented in Appendix 1 demonstrate that newer versions of a drug are 7 to 20 percent more expensive on average. This suggests that the CPI constraint described above would distort firms' behavior because the optimal price for a particular drug appears to be increasing over time.[25]

IV.B. New Drug Introductions

Our empirical strategy to test whether Medicaid regulations increase the number of versions of a drug that are introduced is as follows. Using our Medicaid utilization data, we can determine when new versions of a drug enter the market. We count the number of new NDCs introduced between 1997 and 2002. We focus on our earlier sample because we are interested in the entry of new versions of a drug, which may take several years.

We control for factors that are likely to influence the rate at which new versions of a drug are introduced. The first *log(Sales97)*, is simply equal to the log of total U. S. revenues for the drug in 1997. Because of the fixed costs of introducing a new version, as the total revenues for a drug increase so do the potential gains to releasing new versions. Likewise, we control for generic competition at the start of the period (1997), as this will affect growth of revenue in the future. Note that we have no ex ante prediction on the sign of this variable. It could be negative because of market share (revenue) erosion, or it could be positive, as the brand attempts to differentiate itself from generic competition. We also assume branded manufacturers correctly anticipate generic entry

25. As shown in the earlier version of this paper [Duggan and Scott Morton 2004], an even more rapid increase occurs across drugs. For example, drugs released in 2001 or 2002 are more than twice as expensive on average as those released from 1995 to 1997.

and subsequent revenue loss, and so we include a generic competition indicator at the end of the period (2002). Revenue loss in 2002 clearly affects the returns from launching a new product in 1997. The age of the drug is an additional variable that potentially measures the returns from proliferation, as it embodies the technology of the drug. Last, future returns are also affected by competition from therapeutic substitutes, so we include the same measure used in the price regression above. We control for the number of NDCs in 1997 (*NumNDC*), as this may capture underlying demand for different versions and the costs of developing those versions when estimating specifications of the following form:

$$(2)\quad NewNDC_{j,97-02} = \phi_0 + \phi_1 MMS_{j,97} + \phi_2 GenComp_{j,97} + \phi_3 Subst_{j,97} + \phi_4 Years_{j,97} + \phi_5 log(Sales_{j,97}) + \phi_6 NumNDC_{j,97} + \phi_8 GenComp_{j,02} + \sum_{k=1}^{14} \phi_k^* I(Class_{j,97} = k).$$

In this regression the variable *NewNDC* is simply equal to the number of NDCs with more than 0.1 percent of a drug's sales in 2002 and that had no utilization in 1997. We run Poisson regressions in this case given that the values for this variable are all integers and range from 0 to 6, though the estimates from a set of negative binomial regressions are almost identical.

The first column of Table V summarizes the results. As predicted, the estimate for the Medicaid share is positive, though it is not statistically significant. The estimate for the *Log(Sales)* coefficient is significantly positive, implying that large drugs introduce more new versions, while the coefficient estimates for *Years on Market and Generic Competition* are negative and significant for the former only; old drugs introduce fewer versions. This is logical because their cumulative future revenues are smaller and so the higher price they would obtain from a new version is worth less to them. In the second column of the table, therefore, we focus on newer drugs, drugs that were launched six or fewer years before 1997.[26] This leaves us with the half of the sample that has more to lose from a constrained low price because they will not face generic competition as soon, and the technology

26. One can vary the cutoff for number of years between four and twelve, and the estimated coefficient increases as the sample is restricted to more recent drugs.

TABLE V
POISSON MAXIMUM LIKELIHOOD ESTIMATES OF IMPACT OF MEDICAID ON NDC INTRODUCTIONS

	# of New NDCs Introduced from 1997 to 2002			
	(1)	(2)	(3)	(4)
Medicaid market share in 1997	0.445	1.197*		
	(.723)	(.629)		
Medicaid sales 50–99 million			0.072	0.389
			(.347)	(.472)
Medicaid sales 100–199 million			0.331	0.686
			(.424)	(.755)
Medicaid sales 200 million+			0.441	1.234**
			(.565)	(.589)
Log(sales97)	0.520***	.434**	0.415**	0.197
	(.122)	(.178)	(.167)	(.201)
Any generic competition in 1997	0.019	−0.511	−0.014	−0.549
	(.364)	(.702)	(.377)	(.679)
# of substitutes in 1997	−0.021	−0.027	−0.021	−0.025
	(.014)	(.017)	(.014)	(.018)
Years on market in 1997	−0.075***	−0.068	−.077**	−0.090
	(.029)	(.085)	(.031)	(.083)
Any generic competition in 2002	0.108	0.311	0.091	0.373
	(.288)	(.550)	(.289)	(.525)
Number of NDCs in 1997	0.138***	0.103	0.143***	0.146*
	(.038)	(.077)	(.040)	(.080)
# observations	200	105	200	105
Class fixed effects?	Yes	Yes	Yes	Yes
Pseudo R^2	0.165	0.154	0.167	0.162

Table V summarizes the results from Poisson maximum likelihood specifications in which the dependent variable is the number of new versions of a drug introduced between 1997 and 2002. Specifications summarized in columns (1) and (3) include the top 200 drugs (excluding over-the-counter and generic) based on total sales in the United States in 1997. Specifications (2) and (4) include only those drugs from this group that were introduced in 1991 or later. Each specification includes fourteen therapeutic class fixed effects. Huber-White standard errors are reported in parentheses. ***, **, and * represent significance at the 1, 5, and 10 percent levels, respectively.

embodied in their drugs is newer and in higher demand. This set of drugs responds much more strongly to the Medicaid incentives, as we expected. In column (2), Medicaid market share is significantly (at the 10 percent level) positively related to the number of new NDCs introduced.

In the next two columns we list the results from analogous specifications that include three indicator variables (50–99 million, 100–199 million, and 200+ million) for total sales to Medicaid. If there are fixed costs to launching a new version, the

dollar amount of sales to Medicaid should affect the optimal rate of introduction of new versions. These results yield a similar picture, suggesting that new drugs with very high Medicaid sales are significantly more likely to introduce new versions. For example, the positive and significant coefficient on *Medicaid Sales 200 million* + in column (4) implies that, after controlling for total sales, the rate of new drug introduction for drugs with more than $200 million in Medicaid sales is more than twice as high as for those with less than $50 million in Medicaid sales. Similarly, firms producing drugs with Medicaid sales between $100 and 200 million and $40 and 99 million introduce new versions at higher, though statistically insignificant, rates than firms selling otherwise similar drugs with low Medicaid sales. This final set of results therefore suggests that firms respond to the Medicaid CPI constraint, which aims to restrain the growth rate of pharmaceutical spending, by introducing new and more expensive versions of their drugs.

V. Conclusion

The results summarized in this paper demonstrate that prescription drugs sold disproportionately to Medicaid recipients have significantly higher prices than otherwise similar drugs. Because the Medicaid price is based on prices paid in the private sector, firms have a strong incentive to charge private patients more for their drugs than they otherwise would. To the extent that this happens, some private patients who would otherwise purchase these drugs are priced out of the market. Our second set of findings, that Medicaid's CPI constraint creates additional product proliferation beyond what would exist in the absence of this constraint, is somewhat more subtle. The regulation only limits Medicaid price increases for existing products, so a firm can introduce a new product to obtain freedom to raise prices. Because of the fixed cost of introducing a new version, this strategy is only attractive for drugs with large sales to Medicaid and drugs expected to continue to have large sales to Medicaid. Our evidence indicates that the manufacturers of these drugs find the benefit of a new version outweighs the fixed cost required; these drugs proliferate more new NDCs than would otherwise be expected during our study period.

The combination of these two findings leads us to conclude that the Medicaid procurement regulations distort market out-

comes. More generally, our results suggest that governments face a trade-off when determining how optimally to set prices for any good or service that they procure. The benefits of using private sector prices is that governments have a difficult time "getting prices right," and price has both important distributional consequences and substantial effects on innovation incentives. But a potentially important cost is the distortion of equilibrium outcomes in the private market, with this effect increasing with the government's share of the market.

Medicaid's price-setting rules may substantially lower allocative efficiency for certain drug treatments, as some low valuation Medicaid recipients take the drug while other high-valuation consumers without Medicaid coverage do not. Additionally, non-Medicaid consumers may purchase health insurance to reduce their medical expenditure, thus leading to further increases in prices.[27] On the bright side, given its effect on pharmaceutical prices, Medicaid's policies will increase a firm's potential profits from entry and thus could lead to more innovation for drugs that differentially benefit low-income individuals. More work that examines the effect of Medicaid and other government programs on innovation in this sector is clearly warranted.[28]

Our results take on additional policy significance when one considers that government involvement in the pharmaceutical industry will soon expand by a great deal due to the Medicare prescription drug benefit, which must necessarily have procurement policies also. At the same time, the impact of the Medicaid rules themselves will decline, as Medicaid recipients who are also eligible for Medicare will be shifted to the Medicare drug benefit. While less than 15 percent of Medicaid recipients are dually eligible for Medicare, they account for 60 percent of Medicaid prescription drug spending.[29] Estimating the incidence of the Medicare prescription drug benefit represents an important area for future research.

27. Alternatively, individuals could try to become eligible for Medicaid given its increasing value as pharmaceutical prices rise. See Yelowitz [2002] for an estimation of the effect of state-level Medicaid generosity on SSI receipt.

28. See Finkelstein [2004] for an investigation of the effect of vaccine coverage by the federal government.

29. See Newhouse [2004] for a discussion of this legislation and the likely impact of alternative reimbursement rules. Anecdotal evidence suggests that it has already increased pharmaceutical prices [Martinez 2004].

Appendix 1: New Treatments

We estimate the following specification using all NDCs for the top 200 drugs in both 1997 and 2002:

$$(3)\quad log(Price_j) = a + \theta * LaterVersion_j + \sum_{k=1}^{200} \lambda * I(Drug_j = k),$$

and our results are summarized in Table VI. The explanatory variable of interest is $LaterVersion_j$, which takes on a value of zero if version j of a drug enters the market in the year that the drug is first available and one if it enters in a subsequent year. The statistically significant estimate of .108 for θ in the first column of Table VI suggests that newer versions of a drug are approximately 11 percent more expensive than the first version and this result is significant at the 10 percent level. The results for the analogous specification using the 2002 sample of drugs yield slightly smaller estimates for θ, suggesting a 7 percent increase in average price from initial to subsequent versions. Thus, this is suggestive evidence of the optimal price for a drug increasing over time.

Given that new versions of a drug may—on average—have different dosage amounts than earlier ones, an alternative and

TABLE VI
Are New Versions of a Drug More Expensive?

	Log(price per prescription)		Log(price per unit)	
	(1)	(2)	(3)	(4)
Later Version	0.108*	0.071	0.191***	0.203***
	(.057)	(.057)	(.072)	(.064)
# Observations	728	851	728	851
Sample	1997	2002	1997	2002
R^2	0.782	0.865	0.907	0.942
Drug fixed effects?	Yes	Yes	Yes	Yes
Drug-route fixed effects?	No	No	Yes	Yes

Specifications include all NDCs for the top 200 drugs (excluding over-the-counter and generic) in 1997 or 2002 in terms of total U. S. revenues. The dependent variable in specifications (1) and (2) is the log of the average price per prescription for each NDC while in (3) and (4) it is the log of the average price per unit. Later Version is equal to 0 if an NDC enters the market in the year that a drug is first available and 1 if in a subsequent year. We exclude NDCs for which we do not have an entry date, which is less than 10 percent of the sample in both years. The first two specifications include 200 drug fixed effects, while the last two include 287 drug-route fixed effects. Huber-White standard errors are reported in parentheses. ***, **, and * represent significance at the 1, 5, and 10 percent levels, respectively.

perhaps more appropriate measure of a drug's price is the cost per unit. Thus, we next use information in the Medicaid data on total units dispensed in 1997 and in 2002 to calculate this for each NDC. The next two columns of Table VI summarize the results from specifications using this dependent variable. In these regressions we interact the drug fixed effects with indicators for route type (e.g., tablet, liquid, ointment, etc.) given that units for the same drug may not otherwise be comparable. Thus, the number of fixed effects rises to 287 (from 200) in the 1997 sample and 279 in the 2002 sample. The estimates for θ in this specification are approximately twice as large, once again suggesting that newer versions of a drug tend to be more expensive.

Appendix 2: Data

1. Construction of the 1997 and 2002 Samples

We use the IMS National Sales Perspectives (TM) data to identify the top-selling 200 drugs in each year that are neither generic nor available over-the-counter in the relevant year. If a drug's name is the same as its active ingredient, we assume that it is generic. Revenues for different versions of the same drug (e.g., Prozac and Prozac Weekly) are aggregated to one observation.

2. Calculating Medicaid Sales for Each Drug in 1997 and 2002

We use utilization data that are available from the Centers for Medicare and Medicaid Services (CMS) web site at *http://www.cms.hhs.gov/medicaid/drugs/drug5.asp*. These data provide quarterly information for the United States and for each individual state on the number of prescriptions filled for Medicaid recipients and total Medicaid expenditures at the national drug code (NDC) level. These data are then merged to CMS drug product data to determine the name for each NDC along with information about whether the drug is generic, available over-the-counter, and the route type. The most recent version of these data, which are updated periodically, can be found at *http://www.cms.hhs.gov/medicaid/drugs/drug6.asp*. Some NDCs in the utilization data are not available in the drug product data. For these NDCs, first merge with data obtained from previous versions of the Food and Drug Administration's (FDA's) National Drug Code Directory. The most recent version is available at *http://www.fda.gov/cder/ndc/index.htm*. If information on an NDC is not available in

either of these sets of files, then use the abbreviated name in the utilization data themselves to determine the name of the drug. Select only those NDCs that are included in the sample of interest to calculate Medicaid market share in the relevant year. Medicaid revenues for different versions of the same drug are aggregated to one observation. If more than 10 percent of a drug's sales are accounted for by generic or over-the-counter NDCs, then drop the drug from the sample.

3. Construction of Outcome and Explanatory Variables

a. Medicaid market share: total Medicaid expenditures divided by total U. S. revenues. For one drug in 1997 and three in 2002, this exceeds 1.00. This could be because our measure of Medicaid spending does not subtract out Medicaid Rebates, which are not publicly available (we submitted an FOIA request to CMS but were unable to obtain these data), and thus our measures will to some extent overstate Medicaid spending. This could also be because of differences in the timing of reporting, with some pharmaceutical firms using fiscal years rather than calendar years. We test the sensitivity of our results to excluding these drugs or replacing the market share with 1.00 and find little difference.
b. Price per prescription: total Medicaid expenditures divided by total Medicaid prescriptions. This measure is used because the price per prescription is not available in the IMS data. We compared these prices with average prices for all heath care consumers for the top 50 drugs in 2001. These were reported in the National Institute of Health Care Management's (NIHCM's) publication *Prescription Drug Expenditures in 2001: Another Year of Escalating Costs* (available at *http://www.nihcm.org/spending2001.pdf*). An OLS regression of our measure on this one yields a coefficient estimate of 1.036 with a standard error of .024, suggesting that the Medicaid data provide a quite reliable estimate of average prices for all U. S. consumers.
c. Number of NDCs in 1997: the number of versions of a drug accounting for more than 0.1 percent of a drug's Medicaid utilization in each year.
d. Number of new NDCs from 1997–2002: the number of versions of a drug with no utilization in 1997 and that

accounted for more than 0.1 percent of a drug's Medicaid utilization in 2002.

e. Therapeutic category: there are fourteen of these in the IMS data. The two most common ones in the top 200 samples are drugs used to treat mental disorders and heart conditions. Within each of these major categories there is an average of eighteen subcategories.

f. The number of substitutes: this is equal to the number of ingredients with nonzero sales in the relevant year within the drug's therapeutic subcategory (of which there are 249 in the data). If a drug appears in multiple subcategories, then assign it to the one with the largest sales.

g. Any generic competition: this is set equal to one if there is any NDC in the Medicaid data with the same name as the drug's ingredient name. In cases where the two names are very similar but not identical, this was hand-checked. If a drug has multiple ingredients and any of them are generic, then the drug is coded as facing generic competition.

h. Years on the market: the sample year (1997 or 2002) minus the year that the drug entered the market (included in the IMS data). If there are multiple entry years for the same drug, then use the earliest one.

4. *Steps in Calculating the Instrument for Medicaid Market Share*

a. Use the *Med Ad News* list of top 200 drugs in each year to determine for which conditions each drug is indicated.

b. Match this information to ICD-9 codes using the Centers for Disease Control's listing at *ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/ICD-9/ucod.txt*.

c. Use data from the 1997 and 2002 versions of the National Ambulatory Medical Care Survey. These data include detailed information on a "sample of visits to non-federally employed office-based physicians who are primarily engaged in direct patient care." Aggregating across these two surveys, there are 53,453 physician visits.

d. Using the primary diagnosis ICD-9 code along with the information on each patient's primary expected payer, determine what fraction of patients are on Medicaid for each ICD-9 code.

e. Aggregate over all ICD-9 codes for which each drug is indicated to determine the fraction on Medicaid. This predicted Medicaid fraction is used in the IV analyses sum-

marized in Table IV. Drop drugs from the sample that have fewer than twenty patients in the NAMCS. This leads us to exclude 13 drugs from the 1997 sample and 23 from the 2002 sample.

5. Construction of the Sample with the Top-Selling Drugs in Each Subcategory

For each of the 249 therapeutic subcategories, determine which drug is the top-seller in each year. If there are no drugs with nonzero sales in a year, then exclude this subcategory. If there are multiple versions of the same drug, then aggregate to one observation. Use the same algorithm as described in (2) above to calculate Medicaid sales for these drugs. If more than 10 percent of a drug's sales are accounted for by generic or over-the-counter NDCs then exclude this subcategory.

6. Construction of the Generic Sample for Table III

Select the top-selling drugs in which the ingredient name and drug name are identical in the IMS data. Collapse multiple versions of the same drug to one observation. Use the same algorithm as described in (2) above to calculate Medicaid sales for these drugs. Drop drugs for which more than 10 percent of sales are in over-the-counter versions or in brand (nongeneric) versions.

UNIVERSITY OF MARYLAND AND NATIONAL BUREAU OF ECONOMIC RESEARCH
YALE UNIVERSITY SCHOOL OF MANAGEMENT AND NATIONAL BUREAU OF ECONOMIC RESEARCH

REFERENCES

Aldridge, Christopher, and Arnold Doyle, "AIDS Drug Assistance Programs—Getting the Best Price?" mimeo, National AIDS Drug Assistance Program Monitoring Project, April 2002.

Averch, Harvey, and Leland Johnson, "Behavior of the Firm under Regulatory Constraint," *American Economic Review,* LII (1962), 1052–1069.

Borrell, Joan-Ramon, "Pharmaceutical Price Regulation—A Study on the Impact of the Rate-of-Return Regulation in the UK," *PharmacoEconomics,* XV (1999), 291–303.

Currie, Janet, and Jeffrey Grogger, "Medicaid Expansions and Welfare Contractions: Offsetting Effects on Prenatal Care and Infant Health?" *Journal of Health Economics,* XXI (2002), 313–335.

Cutler, David, and Jonathan Gruber, "Does Public Insurance Crowd Out Private Insurance?" *Quarterly Journal of Economics,* CXI (1996), 391–430.

Duggan, Mark, "Do New Prescription Drugs Pay for Themselves? The Case of Second-Generation Antipsychotics," *Journal of Health Economics,* XXIV (2005), 1–31.

Duggan, Mark, and William Evans, "Estimating the Impact of Medical Innovation: The Case of HIV Antiretroviral Treatments," NBER Working Paper No. 11109, Cambridge, MA, National Bureau of Economic Research, 2005.

Duggan, Mark, and Fiona Scott Morton, "The Distortionary Effects of Government Procurement: Evidence from Medicaid Prescription Drug Purchasing," NBER Working Paper No. 10930, Cambridge, MA, National Bureau of Economic Research, 2004.
Epstein, Arnold, and Joseph Newhouse, "The Impact of Medicaid Expansions on Early Prenatal Care and Health Outcomes," *Health Care Financing Review,* XIX (1998), 85–99.
Finkelstein, Amy, "Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry," *Quarterly Journal of Economics,* CXIX (2004), 527–564.
Frank, Richard G., and David S. Salkever, "Pricing, Patent Loss and the Market for Pharmaceuticals," *Southern Economic Journal,* LIX (1992), 165–179.
Frank, Richard G., and David S. Salkever, "Generic Entry and the Pricing of Pharmaceuticals," *Journal of Economics and Management Strategy,* VI (1997), 75–90.
Gruber, Jonathan, "Medicaid," in R. Moffitt, ed., *Means-Tested Transfer Programs in the United States* (Chicago, IL: University of Chicago Press, 2003).
Gruber, Jonathan, and Aaron Yelowitz, "Public Health Insurance and Private Saving," *Journal of Political Economy,* CVII (1999), 1249–1274.
Holmstrom, Bengt, and Paul Milgrom, "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization,* VII (1991), 24–52.
Ikegami, Naoki, Shunya Ikeda, and Hiroki Kawai, "Why Medical Care Costs in Japan Have Increased Despite Declining Prices for Pharmaceuticals," *Pharmoeconomics,* XIV (1998), 97–105.
Kaiser Commission on Medicaid and the Uninsured, "Medicaid Outpatient Prescription Drug Benefits: Findings from a National Survey and Selected Case Study Highlights," Menlo Park: Kaiser Family Foundation, 2001.
——, "Medicaid Enrollment and Spending Trends," Menlo Park: Kaiser Family Foundation, 2005.
Laffont, Jean-Jacques, and Jean Tirole, *A Theory of Incentives in Procurement and Regulation* (Cambridge, MA: MIT Press, 1993).
Martinez, Barbara, "Drug-Price Surge May Erode Savings from Medicare Card," *Wall Street Journal,* March 24, 2004, page B1.
National Institute for Health Care Management, "Prescription Drug Expenditures in 2001: Another Year of Escalating Costs," mimeo, 2002.
Newhouse, Joseph, "How Much Should Medicare Pay for Prescription Drugs?" *Health Affairs,* XXIII (2004), 89–102.
Olson, Mary, "Substitution in Regulatory Agencies: FDA Enforcement Alternatives," *Journal of Law, Economics, and Organization,* XII (1996), 376–407.
Parker, David, "The Performance of BAA Before and After Privatization—A DEA Study," *Journal of Transport Economics and Policy,* XXXIII (1999), 133–145.
Scott Morton, Fiona, "The Strategic Response by Pharmaceutical Firms to the Medicaid Most-Favored-Customer Rules," *RAND Journal of Economics,* XXVIII (1997), 269–290.
Tirole, Jean, *The Theory of Industrial Organization* (Cambridge, MA: MIT Press, 1991).
Yelowitz, Aaron, "The Medicaid Notch, Labor Supply, and Welfare Participation: Evidence from Eligibility Expansions," *Quarterly Journal of Economics,* CV (1995), 909–940.
——, "Why Did the SSI Disabled Program Grow So Much? Disentangling the Effect of Medicaid," *Journal of Health Economics,* XVII (1998), 321–349.