Market Size in Innovation: Theory and Evidence From the Pharmaceutical Industry*

Daron Acemoglu **Joshua Linn**
MIT **MIT**

January 2004.

Abstract

This paper investigates the effect of (potential) market size on entry of new drugs and pharmaceutical innovation. Focusing on exogenous changes driven by U.S. demographic trends, we find that a 1 percent increase in the potential market size for a drug category leads to approximately a 5 percent increase in the number of new non-generic drugs. This response is generally robust to controlling for a variety of non-profit factors, pre-existing trends, and changes in the technology of pharmaceutical research.

Keywords: Economic Growth, Entry, Innovation, Pharmaceuticals, Profit Incentives, R&D.
JEL Classification: O31, O33, L65.

*We thank David Autor, Ernie Berndt, Olivier Blanchard, Alfonso Gambardella, Amy Finkelstein, Stan Finkelstein, Jerry Hausman, Paul Joskow, Frank Lichtenberg, Lawrence Katz, two anonymous referees, and participants at Chicago, MIT, the NBER Summer Institute and Wharton for useful comments. We are especially grateful to Frank Lichtenberg for generously sharing the Food and Drug Administration drug approval and the Computer Retrieval of Information on Scientific Projects data with us. This research is partially funded by the National Science Foundation and the Russell Sage Foundation.

1 Introduction

This paper constructs a simple model linking innovation rates to current and future market size, and provides evidence from the pharmaceutical industry to support this hypothesis. Our empirical work, which exploits changes in the market size for different drug categories driven by U.S. demographic trends, finds economically significant and relatively robust effects of market size on innovation.

Although many historical accounts of important innovations focus on the autonomous progress of science and on major breakthroughs that take place as scientists build on each other's work, economists typically emphasize profit incentives and the size of the target market. For example, in his seminal study, *Invention and Economic Growth,* Schmookler argued that: "...invention is largely an economic activity which, like other economic activities, is pursued for gain" [1966, p. 206]. To emphasize the role of market size, Schmookler entitled two of his chapters "The amount of invention is governed by the extent of the market."

The role of profit incentives and market size in innovation is also important both for the recent endogenous technological change models, which make profit incentives the central driving force of the pace of aggregate technological progress [e.g., Aghion and Howitt, 1992, Grossman and Helpman, 1991, Romer, 1990], and for the induced innovation and directed technical change literatures, which investigate the influence of profit incentives on the types and biases of new technologies [see, for example, Kennedy, 1964, Drandkis and Phelps, 1965, Samuelson, 1965, Hayami and Ruttan, 1970, and Acemoglu, 1998, 2002, and 2003]. A recent series of papers by Kremer, for example [2002], also build on the notion that pharmaceutical research is driven by market size and argue that there is generally insufficient research to develop cures for third-world diseases such as malaria.

In this paper, we investigate the effect of market size on entry of new drugs and pharmaceutical innovation. A major difficulty in any investigation of the impact of market size on innovation is the endogeneity of market size—better products will have larger markets. Our strategy to overcome this problem is to exploit variations in market size driven by U.S. demographic changes, which should be exogenous to other, for example scientific, determinants of innovation and entry of new drugs.[1] To create potential market size, we construct age profiles of users for each drug category at a point in time, and then compute the implied market size from aggregate demographic and income changes given these (time-invariant) age profiles.[2] We

[1] For many drugs non-U.S. markets may also be relevant. Nevertheless, the U.S. market is disproportionately important, constituting about 40 percent of the world market [IMS, 2000]. Below we report results using changes in OECD market size as well as U.S. market size.

[2] Loosely speaking, "market size" corresponds to the number of users times their marginal willingness to

measure entry and innovation using the Food and Drug Administration's (FDA) approval of new drugs.[3]

Our results show that there is an economically and statistically significant response of the entry of new drugs to market size. For example, a 1 percent increase in the size of the potential market for a drug category leads to a 7-10 percent increase in the total number of new drugs. Much of this response comes from the entry of generics, which are drugs that are identical or bioequivalent to an existing drug no longer under patent protection. More important, there is also a statistically significant response of the entry of non-generic drugs, which more closely correspond to "innovation": a 1 percent increase in potential market size leads to a 5 percent increase in the number of new non-generic drugs in the market. Interestingly, while generics respond to current market size, we find that non-generics respond to five-year leads of market size. The response of non-generic entry to anticipated changes in market size in the near future is consistent with the predictions of the theoretical model we use to motivate our empirical investigation.

The effect of market size on the entry of new drugs is generally robust. We obtain similar results when we use different measures of market size, when we exploit changes in OECD market size, and when we control for a variety of non-profit factors, pre-existing trends, and advances in biotechnology.

There are a number of other studies related to our work. First, Schmookler [1966] documents a correlation between sales and innovation, and argues that the causality ran largely from the former to the latter. The classic study by Griliches [1957] on the spread of hybrid seed corn in U.S. agriculture also provides evidence consistent with the view that technological change and technology adoption are closely linked to profitability and market size. Pakes and Schankerman [1984] investigate this issue using a more structural approach, linking R&D intensity at the industry level to factor demands and to growth of output. In more recent research, Scott Morton [1999] and Reiffen and Ward [2004] study the decision of firms to introduce a generic drug and find a positive relationship between entry and expected revenues in the target market. None of these studies exploit a potentially exogenous source of variation in market size, however.

Second, some recent research has investigated the response of innovation to changes in energy prices. Most notably, Newell, Jaffee and Stavins [1999] show that between 1960 and 1980, the typical air-conditioner sold at Sears became significantly cheaper, but not much more

pay. Therefore, market size can increase both because the number of users increases and because their marginal willingness to pay changes. We focus on changes driven by demographics to isolate exogenous changes in market size.

[3] These data were previously used by Lichtenberg and Virahbak [2002], who obtained them under the Freedom of Information Act. We thank Frank Lichtenberg for sharing these data with us.

energy-efficient. On the other hand, between 1980 and 1990, there was little change in costs, but air-conditioners became much more energy-efficient, which, they argue, was a response to higher energy prices. In a related study, Popp [2002] finds a strong correlation between aggregate patents and energy prices.

Third, there is substantial research focusing on innovation in the pharmaceutical industry. Henderson and Cockburn [1996], Cockburn and Henderson [2001], and Danzon, Nichelson and Sousa Pereira [2003] study the determinants of success in clinical trials, focusing mainly on firm and project size. Galambos and Sturchio [1998], Cockburn, Henderson and Stern [1999], Gambardella [2000], Malerba and Orsenigo [2000], and Ling, Berndt and Frank [2003] discuss various aspects of the recent technological developments in the pharmaceutical industry.

Most closely related to this study are Lichtenberg and Waldfogel [2003], Finkelstein [2003], and Cerda [2003]. Lichtenberg and Waldfogel show that following the Orphan Drug Act there were larger declines in mortality among individuals with rare diseases (compared to other diseases), and argue that this is related to the incentives created by the Act to develop drugs for these rare diseases. Finkelstein exploits three different policy changes affecting the profitability of developing new vaccines against 6 infectious diseases: the 1991 Center for Disease Control recommendation that all infants be vaccinated against hepatitis B, the 1993 decision of Medicare to cover the costs of influenza vaccinations, and the 1986 introduction of funds to insure vaccine manufactures against product liability lawsuits for certain kinds of vaccines. She finds that increases in vaccine profitability resulting from these policy changes are associated with a significant increase in the number of clinical trials to develop new vaccines against the relevant diseases.[4] Cerda's [2003] Ph.D. dissertation at Chicago is an independent study of the effect of demographics on innovation in the pharmaceutical sector. Although Cerda uses a somewhat different empirical methodology, he reaches similar conclusions to our study.

The rest of the paper is organized as follows. We outline a simple model linking innovation to market size in the next section. Section 3 briefly explains our empirical strategy, and Section 4 describes the basic data sources and the construction of the key variables. Section 5 presents the empirical results and a variety of robustness checks. Section 6 contains some concluding remarks, while the Appendix gives further data details.

[4]Lichtenberg [2003] also presents evidence suggesting that the types of new drugs changed towards drugs more useful for the elderly after Medicare was established.

2 Theory

We now outline a simple model that illustrates the impact of market size on innovation. The economy consists of a set I of infinitely-lived individuals. Time is continuous $t \in [0, \infty)$. There are two types of goods in this economy. First, a basic good, y, which can be consumed or used for the production of other goods, or for research expenditure. Individual i has an exogenously given endowment $y_i(t)$ at time t. Second, there are J drugs, x_1, ..., x_J, each with a potentially time-varying "quality", $q_1(t)$, ..., $q_J(t)$. Each individual demands only one type of drug. Hence, we partition the set I of individuals into J disjoint groups, G_1,...,G_J with $G_1 \cup G_2 \cup ... \cup G_J = I$, such that if $i \in G_j$, then individual i demands drug j. More specifically, if $i \in G_j$, then his preferences are given by:

$$\int_0^\infty \exp(-rt) \left[c_i(t)^{1-\gamma} \left(q_j(t) x_{ji}(t) \right)^\gamma \right] dt, \tag{1}$$

where r is the discount rate of the consumers (also the interest rate in the economy), $\gamma \in (0,1)$, $c_i(t)$ is the consumption of individual i of the basic good at time t, and $x_{ji}(t)$ is the consumption of individual i of drug j. This Cobb-Douglas functional form and the assumption that each individual only consumes one type of drug are for simplicity and do not affect the main results.[5]

Normalizing the price of the basic good to 1 in all periods, and denoting the price of drug j at time t by $p_j(t)$, the demand of individual $i \in I$ for drug j is $x_{ij}(t) = \gamma y_i(t) / p_j(t)$ if $i \in G_j$, and $x_{ij}(t) = 0$ if $i \notin G_j$. Summing across individuals, total demand for drug j is:

$$X_j(t) = \frac{\gamma Y_j(t)}{p_j(t)}. \tag{2}$$

where $Y_j(t) \equiv \sum_{i \in G_j} y_i(t)$ is the total income of the group of individuals consuming drug j.

At any point in time, there is one firm with the best-practice technology for producing each type of drug, and it can produce one unit of this drug with quality $q_j(t)$ using one unit of the basic good. If there is an innovation for drug line j currently with quality $q_j(t)$, this leads to the discovery of a new drug of quality $\lambda q_j(t)$ where $\lambda > 1$. For the purposes of the model, we think that any new innovation is approved (for example by the FDA) and can be sold to consumers immediately (and is under patent protection indefinitely).

There is free entry into R&D and each firm has access to an R&D technology that generates a flow rate δ_j of innovation for every dollar spent for research on drug j. So if R&D expenditure at time t is $z_j(t)$, the flow rate of innovation (and of entry of new drugs) for drug j is:

$$n_j(t) = \delta_j z_j(t). \tag{3}$$

[5] The Cobb-Douglas assumption implies that the share of income spent on drugs is constant. This implication can be easily relaxed by considering a utility function with an elasticity of substitution different from 1, as in the factor market models with directed technical change [see, for example, Acemoglu, 1998, 2002].

Differences in δ_j's introduce the possibility that technological progress is scientifically more difficult in some lines than others.

A key feature of this R&D technology for our focus is that research is *directed* in the sense that firms can devote their R&D to developing particular types of drugs. The pharmaceutical industry, especially in the recent past, is a prime example of an industry where companies with fairly sophisticated R&D divisions or specialized R&D firms can undertake research for specific drug lines [e.g., Gambardella, 2000, Malerba and Orsenigo, 2000].[6]

The demand curves in (2) have an elasticity equal to 1, so an unconstrained monopolist would charge an arbitrarily high price. However, the firm with the best drug in line j is competing with the next best drug in that line. An arbitrarily high price would allow the next best firm to capture the entire market. Therefore, the firm with the best drug sets a *limit* price to exclude the next best firm—i.e., to ensure that consumers prefer its product rather than the next best drug supplied at the lowest possible price equal to marginal cost, which is 1. If a consumer buys from the best-practice firm with quality $q_j(t)$ and price $p_j(t)$ and chooses her optimal consumption as given by (2), her instantaneous utility at time t is $(q_j(t))^{\gamma}(1-\gamma)^{1-\gamma}\gamma^{\gamma}(p_j(t))^{-\gamma}y_i(t)$; and if she purchases from the next best firm, which, by definition, has quality $q_j(t)/\lambda$ at price equal to marginal cost, 1, she obtains utility $\lambda^{-\gamma}(q_j(t))^{\gamma}(1-\gamma)^{1-\gamma}\gamma^{\gamma}y_i(t)$. The limit price, which equalizes these two expressions, is:

$$p_j(t) = \lambda \text{ for all } j \text{ and } t. \tag{4}$$

The profits of the firm with the best product of quality $q_j(t)$ in line j at time t are:

$$\pi_j(q_j(t)) = (\lambda - 1)\gamma Y_j(t). \tag{5}$$

Here $\lambda\gamma Y_j(t)$ corresponds to the market size (total sales) for drug j. Notice that profits of drug companies are independent from quality, $q_j(t)$, which is a feature of the Cobb-Douglas utility.

Firms are forward-looking, and discount future profits at the rate r. The discounted value of profits for firms can be written by a standard dynamic programming recursion. $V_j(t \mid q_j)$, the value of a firm that owns the most advanced drug of quality q_j in line j at time t, is:[7]

$$rV_j(t \mid q_j) - \dot{V}_j(t \mid q_j) = \pi_j(q_j(t)) - \delta_j z_j(t) V_j(t \mid q_j), \tag{6}$$

[6] Naturally, there exist examples of research directed at a specific drug type leading to the discovery of a different product, such as the well-known example of Viagra, which resulted from research on hypertension and angina, and was partly accidentally discovered from the detection of side effects in a clinical study [see, e.g., Kling, 1998]. The working paper version, Acemoglu and Linn [2003] shows that the results here generalize even when there is a large component of random R&D, whereby research directed at drug j can result in the discovery of other drugs.

[7] Throughout, we assume that the relevant transversality conditions hold and discounted values are finite.

where $\pi_j(q_j(t))$ is the flow profits given by (5), and $z_j(t)$ is R&D effort at time t on this line by other firms.[8] Intuitively, the value of owning the best technology in line j, $rV_j(t \mid q_j)$, is equal to the flow profits, $\pi_j(q_j(t))$, plus the potential appreciation of the value, $\dot{V}_j(t \mid q_j)$, and takes into account that at the flow rate $n_j(t) = \delta_j z_j(t)$ there will be a new innovation, causing the current firm to lose its leading position and to make zero profits thereafter.

Free entry into R&D to develop better quality drugs implies zero profits, i.e.:

$$\text{if } z_j(t) > 0, \text{ then } \delta_j V_j(t \mid q_j) = 1 \text{ for all } j \text{ and } t. \tag{7}$$

(and if $z_j(t) = 0$, $\delta_j V_j(t \mid q_j) \leq 1$ and there will be no equilibrium R&D for this drug).

An equilibrium in this economy is given by sequences of prices $p_j(t)|_{j=1,..J}$ that satisfy (4), consumer demands for drugs $x_i(t)|_{i \in I}$ that satisfy (2) and R&D levels $z_j(t)|_{j=1,..J}$ that satisfy (7) with $V_j(\cdot)$ given by (6).

An equilibrium is straightforward to characterize. Differentiating equation (7) with respect to time implies $\dot{V}_j(t \mid q_j) = 0$ for all j and t, as long as $z_j(t) > 0$. Substituting this equation and (7) into (6) yields the levels of R&D effort in the unique equilibrium as:

$$z_j(t) = \max\left\{\frac{\delta_j(\lambda - 1)\gamma Y_j(t) - r}{\delta_j}; 0\right\} \text{ for all } j \text{ and } t. \tag{8}$$

Equation (8) highlights the market size effect in innovation: the greater is $Y_j(t)$, the market size for a particular drug, the more profitable it is to be the supplier of that drug, and consequently, there will be greater research effort to acquire this position. In addition, a higher productivity of R&D as captured by δ_j also increases R&D, and a higher interest rate reduces R&D since current R&D expenditures are rewarded by future revenues.

Another important implication of this equation is that there are no transitional dynamics. At any point in time, R&D for a particular drug line is determined by the current market size—past and anticipated future market sizes do not affect current research effort. This is an implication of the linear R&D technology, which ensures that whenever there are profit opportunities, there will immediately be sufficient R&D to arbitrage them, ensuring $\dot{V}_j(t \mid q_j) = 0$. The intuition for the lack of response to anticipated changes in future market size highlights an important effect in quality ladder models of technological progress: firms would like to own the best-practice product at the time the market size actually becomes larger. Investing in R&D far in advance of the increase in market size is not profitable, since another firm would improve over this innovation by the time the larger market size materializes. In fact, with the

[8] Because of the standard replacement effect first emphasized by Arrow [1963], the firm with the best technology does not undertake any R&D itself (see, for example, Aghion and Howitt [1992]).

linear model here, z_j can change discontinuously, so investing even a little bit in advance of the actual increase in the size of the market is not profitable.

Combining equations (3) and (8) gives entry of new drugs as:

$$n_j(t) = \max\left\{\delta_j(\lambda - 1)\gamma Y_j(t) - r; 0\right\}. \tag{9}$$

This equation relates innovation or entry of new products to market size (total expenditure of consumers in this line of drug). It also encompasses the alternative view of the determinants of innovation, discussed in the Introduction, that the cross-drug distribution of R&D is determined by technological research opportunities or perhaps by other non-profit related motives. If there are large and potentially time-varying differences in δ_j's, then these may be the primary factor determining variation in R&D across drug lines, and market size may have only a small effect. Whether or not this is so is an empirical question.

The working paper version of our paper, Acemoglu and Linn [2003], presented a number of generalizations of this framework. First and most importantly, we modified the R&D technology captured in equation (3) to allow for within-period decreasing returns, so that

$$n_j(t) = \delta_j z_j(t)\,\phi(z_j(t)),$$

where $\phi'(z) \leq 0$ (the model studied above is the special case with $\phi'(z) \equiv 0$). Most of the results here generalize, but the model also implies a potential response to anticipated changes in future market size. In particular, let us assume that $Y_j(t) = Y_j$ for all t. Then it is straightforward to show that steady-state R&D will be given by

$$z_j^S = \max\left\{\frac{\left(\delta_j\phi\left(z_j^S\right)(\lambda - 1)\gamma Y_j - r\right)}{\delta_j\phi\left(z_j^S\right)}; 0\right\},$$

which is similar to (8). If there is an unanticipated change in Y_j, there continues to be no transitional dynamics (i.e., z_j immediately jumps to its new steady-state value). But it can be shown that if there is an *anticipated* increase in market size in the future, there will be entry of new drugs in advance of the actual increase. Nevertheless, the same forces emphasized here imply that investing in R&D too far in advance would not be profitable because another firm is likely to innovate further before the actual increase in market size materializes. In terms of our empirical work, even if demographic changes are anticipated 20 or 30 years in advance, we may expect significant innovation responses much later, perhaps 5 or 10 years in advance.

Second, we also extended this model to incorporate entry of both generic and non-generic drugs and showed that market size has a positive effect on entry of both types of drugs, and that, under plausible circumstances, there should be a larger effect of market size on generic drug entry.

3 Empirical Strategy

3.1 Empirical Specification and Estimation Issues

As $r \to 0$, we can take logs on both sides of equation (9) to obtain:

$$\log n_j(t) = \text{constant} + \log \delta_j + \log m_j(t),$$

where $m_j(t) \equiv \lambda\gamma Y_j(t)$ is the market size for drug line j at time t. We measure entry of new drugs (or innovation), $n_j(t)$, as new drug approvals by the FDA in broad drug categories as described below. This measure, denoted by N_{ct} for drug category c at time t, includes entry of generic drugs. Although generic drugs do not correspond to "innovation", their entry is driven by the same profit incentives as innovation. After presenting results using all drug approvals, much of our analysis focuses on the relationship between market size and entry of non-generics. Throughout, instead of actual market size, $m_j(t)$, we use potential market size driven by demographic changes, which we denote by M_{ct}, and discuss its construction below.

Adding other potential determinants, time effects and an error term capturing other unobserved influences, and allowing the coefficient of $\log M_{ct}$ to differ from 1 as it would with more general preferences than Cobb-Douglas, we arrive at an estimating equation of the form:

$$\log N_{ct} = \alpha \cdot \log M_{ct} + X'_{ct} \cdot \beta + \zeta_c + \mu_t + \varepsilon_{ct}, \tag{10}$$

where N_{ct} is the number of new drugs in category c in time period t, M_{ct} is potential market size, X'_{ct} is a vector of controls, including a constant, ζ_c's are a full set of category fixed effects that correspond to the δ_j terms above, μ_t's are a full set of time effects capturing any common time component, and finally, ε_{ct} is a random disturbance term, capturing all omitted influences. The specification with the dependent variable in logarithm is useful, since it ensures that drug category fixed effects and time effects have proportional impacts on entry of new drugs.

One problem with equation (10) is that N_{ct} is a count variable (number of new drugs), so it can equal 0, making it impossible to estimate (10). So instead we consider the Poisson model (see Hausman, Hall, and Griliches, 1984, and Wooldridge, 1999, 2002):

$$N_{ct} = \exp(\alpha \cdot \log M_{ct} + X'_{ct} \cdot \beta + \zeta_c + \mu_t) + \varepsilon_{ct}, \tag{11}$$

which can be obtained from (9) by adding time effects and other covariates multiplicatively and a random disturbance term, ε_{ct}, additively, and then rearranging.

The estimation of (11) would lead to biased estimates, however, since the nonlinearity in (11) makes it impossible to estimate the fixed effects, the ζ_c's, consistently. To deal with this

problem, we follow Hausman, Hall, and Griliches (1984), and transform (11) to obtain:

$$S_{ct} = \frac{\exp(\alpha \cdot \log M_{ct} + X'_{ct} \cdot \beta + \mu_t)}{\sum_{\tau=1}^{T} \exp(\alpha \cdot \log M_{c\tau} + X'_{c\tau} \cdot \beta + \mu_\tau)} + \varepsilon_{ct}, \tag{12}$$

where $S_{ct} = N_{ct} / \sum_{\tau=1}^{T} N_{c\tau}$ is the number of drugs approved in category c at time t, divided by the total number of drugs approved in category c, and T is the total number of time periods in the sample. This transformation removes the drug category dummies, and the coefficient of interest, α, can be estimated consistently. We estimate this equation using nonlinear least squares (NLLS). Woodridge [1999] shows that NLLS estimation strategy has good consistency properties, even when the true model is not Poisson.

We also estimate (10) with OLS following a procedure introduced by Pakes and Griliches [1980], whereby the left-hand side variable is changed to $\log \tilde{N}_{ct}$ where $\tilde{N}_{ct} = N_{ct}$ if $N_{ct} \geq 1$ and $\tilde{N}_{ct} = 1$ if $N_{ct} = 0$, and a dummy that equals 1 when $N_{ct} = 0$ is added to the right-hand side. This procedure is simple and flexible, but the estimates are biased since the dummy variable for $N_{ct} = 0$ is endogenous.

In addition, we estimate equations with leads and lags of $\log M_{ct}$ to determine whether there are significant delays and anticipation effects. Such effects are possible, since, as reported by DiMasi et al. [1991], it can take as long as 15 years for a drug to enter the market from the time of initial research. Furthermore, changes in demographics can be anticipated a long time in advance, so, as discussed in Section 2, drug approvals may respond to anticipated future market sizes.[9]

3.2 Potential Market Size and Identification

Throughout, we exploit the potentially exogenous component of market size driven by demographic trends, combined with differences in the age profiles of expenditure and use for different types of drugs. We obtain the age profiles from micro drug consumption data, and the changes in U.S. demographics from the CPS (Current Population Survey) data. Our (income-based) measure of potential market size is:

$$M_{ct} = \sum_{a} u_{ca} \cdot i_{at}, \tag{13}$$

where i_{at} is the income of individuals in age group a at time t in the United States, and u_{ca} gives the age profile for drug category c. It is computed as the average expenditure share of

[9] An additional issue is that the FDA approval process may be faster for more profitable drugs, and thus potentially for drugs with greater market size. See Dranove and Meltzer [1994]. Our data do not enable us to investigate this issue.

drugs in category c in the total income of those in age group a. This income-based measure corresponds closely to the market size in the theoretical model, which is a combination of the number of consumers and their incomes. We also check the robustness of our results with an alternative population-based measure, calculated using the U.S. population for age group a at time t for i_{at}, and the average number of drugs in category c used per person in age group a for u_{ca}. It is important that the over-time source of variation in both measures is not from changes in individual use, but purely from demographic changes captured by i_{at}—i.e., u_{ca}'s are not time-varying.[10] Consequently, changes in prices and drug quality, which may result from innovations and affect consumption patterns, will not cause over-time variation in M_{ct}. Our baseline measure uses five-year age groups and time periods corresponding to five-year intervals. We also check the robustness of our results using single year age groups and ten-year intervals.

The major threat to the validity of our empirical strategy is from potentially time-varying omitted variables (the drug category fixed effects take out any variable that is not time-varying). Omitted variables related to market size or profit opportunities may induce a bias in the implied magnitudes, but will not lead to spurious positive estimates of the effect of market size (in other words, the presence of such variables is essentially equivalent to mismeasurement of the appropriate market size). More threatening to our identification strategy would be omitted supply-side variables. If our instrument is valid, it will be orthogonal to variation in supply-side determinants of innovation. We attempt to substantiate our identifying assumption further by checking for residual serial correlation and pre-existing trends, and showing the robustness of our estimates to controlling for potential supply-side determinants of innovation and entry.[11]

4 Data and Descriptive Statistics

The demographic data come from the March CPS, 1965-2000. We compute i_{at} in equation (13) for five-year age groups, ranging from 0-4 to 90+. Individual income is constructed by dividing family income equally among the members of the family. For the purposes of the diagrammatic presentation we aggregate the age groups into five broad categories, 0-20, 20-30, 30-50, 50-60 and 60+. This division is motivated by the similarity of income and population movements of the age groups within each of these broad groups.

[10] If preferences have a Cobb-Douglas form as in (1) and are stable, the expenditure measure of u_{ca} should be constant. With non-Cobb-Douglas preferences, changes in prices and drug quality will induce changes in u_{ca}. In this case, by using a time-invariant measure of u_{ca}, we remove this potentially endogenous source of variation.

[11] Another source of endogeneity may be that innovations in certain drug categories extend the lives of the elderly, thus increasing their M_{ct}. Lichtenberg [2002, 2003] provides evidence that new drugs extend lives. This source of endogeneity is not likely to be quantitatively important, however, since the variation resulting from extended lives in response to new drugs is a small fraction of the total variation in M_{ct}. Nevertheless, we also report estimates that instrument M_{ct} with past demographics, purging it from changes in longevity.

Figure I shows population shares, and Figure II shows the corresponding income shares (i.e., income of the corresponding age group divided by total income in that period) for the five broad age groups. To facilitate comparison with Figure III, Figure II starts in 1970. Both figures show a large amount of variation across age groups over time. In particular, it is possible to trace the baby boomers, as the fraction of those in the age bracket 20-30 in the 1970s, and those in the age bracket 30-50 in the 1980s and the 1990s.

The FDA classifies all prescription drugs into 20 major drug categories, which are further subdivided into 159 categories. These categories are based on a combination of therapeutic intent and chemical structure. We drop 4 of the 20 major categories from this classification: Anesthetics, Antidotes, Radiopharmaceuticals and Miscellaneous.[12] We then subdivide some of these categories according to the conditions and diseases that the drugs are used to treat.[13] For example, within the Hematologics major category, we separate Anemia drugs from Anticoagulants because they treat different diseases. We also subdivide broader groups when the age distribution of expenditure is sufficiently heterogeneous. For example, the indications of drugs in Estrogens/Progenstins and Contraceptives overlap somewhat, but the age structure of users is quite different: 20-30 year-olds use Contraceptives most, while 50-60 year-olds use Estrogens/Progestins most. In one case, we combine categories from different major classes, Antifungals and Dermatologics, because the drugs have similar indications and age distributions. The result is a classification system with 33 categories.[14] Appendix Table A1 lists the 33 categories.

Our main data source for drug use is the Medical Expenditure Panel Survey (MEPS), which is a sample of U.S. households over the years 1996-1998. The survey has age and income data for each household member, and covers about 28,000 individuals each year. There is also a list of prescription drugs used by each person (if any), and the amount spent on drugs, which includes copayments and payments by insurance companies and government programs (e.g., Medicaid and worker's compensation).[15] In all, there are about 500,000 medications

[12] We drop the Anesthetics, Radiopharmaceuticals and Miscellaneous categories because most of the items in these categories were not developed for a distinct market. Radiopharmaceuticals are used for diagnostic purposes, and the Miscellaneous category mainly contains surgical and dental tools. The Antidote category is dropped because there are few drugs approved and there is little use of these drugs in the surveys. See the Data Appendix for further details on the contruction of our categories.

[13] Other authors, for example Lichtenberg [2003], have used a more detailed classification system based on diseases. We were unable to construct a comprehensive mapping of the prescription drugs listed in the micro-data surveys to the detailed disease classes. Our classification system relies on the FDA categories, but then subdivides those according to disease and age distribution.

[14] The working paper version, Acemoglu and Linn [2003], used a system with 34 categories based on FDA classification. Results using this alternative classification are reported in the Appendix. Further details on the construction of the 33 categories used here and on our old classification system are available upon request.

[15] Respondents list the pharmacy or medical provider where they obtained the prescription drug, which are then contacted to validate the information and to gather additional information on prescription drug payments.

prescribed. We compute drug expenditure and use by five-year age groups, then divide these by the corresponding income and population numbers from the CPS to construct the income-based and the population-based measures of u_{ca}. Appendix Table A1 reports these numbers aggregated to the five broad age groups used in the figures. This table shows a large amount of variation in the age profiles of expenditure across the 33 drug categories. The elderly spend more on many categories than do younger individuals, but there are numerous exceptions. For example, Antibiotics are used most by individuals in the youngest group, Contraceptives are used most by 20-30 and 30-50 year-olds, and Antivirals are used most by 30-50 year-olds. We construct the measures of potential market size according to equation (13) by combining expenditure data from the MEPS and income data from the CPS.

To investigate the stability of the age profile of users, we supplement the MEPS data with the National Ambulatory Medical Care Survey (NAMCS), which is an annual survey of doctors working in private practice and includes drug use data for the years 1980, 1981, 1985 and 1989-2000. Observations are at the doctor-patient-visit level; there are about 40,000 visits per year. Doctors are selected randomly, surveyed for a week, and patient-visits are then selected randomly from all the visits that week (further details on this survey are given in Acemoglu and Linn [2003]). We use the same classification system with the NAMCS as with the MEPS. Because the NAMCS does not contain expenditure information and its sampling scheme makes it less representative and less reliable than the MEPS, we focus on the MEPS for our main analysis and use the NAMCS only to check the stability of the age profiles of users.

Table I gives correlations between various measures of drug use. The first two rows of Panel A show a high degree of correlation between the NAMCS surveys at various dates, both unweighted or weighted by total use of each category in the survey. These results indicate that the age profiles of use have not changed significantly over the 1980s and the 1990s. The third and fourth rows report mean correlation by drug. These are constructed by computing the within category correlation between the two measures and then averaging it across all categories. This measure, which is more informative for the question of whether or not the age profiles of use for a particular drug has changed, also shows a substantial degree of persistence over time, especially when we look at the weighted correlation in row 4. The difference between the weighted and the unweighted correlations reflects the relatively imprecise estimates of use per person in five-year age brackets for the smaller categories.

Panel B performs the same calculation for the three waves of the MEPS (weighted correlations now use total expenditure in each category as weights), and similarly shows substantial persistence in the age profiles of expenditure. Notably, there is now an even larger difference

between weighted and unweighted mean correlations by drug, presumably because the MEPS, which is a representative sample of the U.S. population, has only a few observations in some of the smaller drug categories. This motivates our focus below on weighted regressions. Finally, Panel C shows that the NAMCS and the MEPS measures, and also expenditure shares and use per person in the MEPS are highly correlated.

The last major data source is a list of FDA new drug approvals. We exclude over-the-counter drugs, the so-called orphan drugs,[16] and drugs that have the same identifying characteristics (i.e., same name, company, and category, or the same FDA approval number). We focus on the time period 1970-2000. Both the quality of the approvals data and the quality of our measures of potential market size deteriorate as we go back in time, because we can only match FDA categories for drugs that are still listed by the FDA, and we are using age profiles from the 1990s. Our approvals dataset for 1970-2000 comprises 5,374 prescription drugs, including both generics and non-generics (see the Appendix). Since 1970 there have been about 2/3 as many approved non-generics as generics.

Figure III shows the share of drug approvals over time to compare with changes in income shares depicted in Figure II. To construct Figure III, we allocate each of the 33 categories to the five broad age group that has the largest expenditure in that category (there are no categories for which the 20-30 group has the largest expenditure, so we only have 4 curves in the figure, see Appendix Table A1). The share of drug approvals is computed as the number of approvals in a given category in each five-year period divided by total approvals in that period.[17] Although this cut of the data uses only a small part of the information that the regression analysis below exploits, a positive association between changes in income shares and changes in drug approvals can be detected by comparing this figure to Figure II . For example, the income share of the 30-50 group increases over the sample, and so does the entry of drugs most used by this group. The shares of income and entry of drugs for those 0-20, on the other hand, show a downward trend, though the decline in the share of new drugs is somewhat less pronounced and later than the decline in income share. We also see that the income share of 50-60 year-olds is constant at first, then declines and finally increases again. The share of drugs most used by this group first increases slightly, and then declines and then increases again. There is little relationship

[16] These drugs treat rare conditions, affecting fewer than 200,000 people. An example is botox, first developed to treat adult dystonia, which causes involuntary muscle contractions. We drop these drugs because we have difficulty matching them consistently, and because they receive special inducements under the Orphan Drug Act.

[17] There are large fluctuations in the total number of approvals, partly because of a number of institutional changes. For example, it was discovered in 1989 that some FDA officials were taking bribes to speed up the approval process for generic drugs. As a result, in the early 1990s the approval process for generics was greatly slowed. See, for example, The Washington Post, August 16, 1989. When we separate our approval data into generics and non-generics, we see a large drop in generics approvals in the early 1990s, but only a small decline for non-generics. We thank Ernie Berndt for suggestions on this issue.

between the share of income and the share of drugs for the 60+ age group, however. Figure IV is similar to Figure III, but is for non-generics; it shows a similar, though noisier, pattern. These patterns are explored in greater detail in the regression analysis below.

5 Results

5.1 Basic Specifications

Table II provides the basic results from the estimation of equation (12) with nonlinear least-squares (NLLS). The top panel is for all approvals. Panel B and C look separately at non-generic and generics. Throughout the paper, the standard errors are corrected for heteroscedasticity using the Huber-White formula. In this table and in our baseline specifications, we use the basic (income-based) measure of $\log M_{ct}$, constructed using expenditure data from the MEPS, and income from the CPS, the time periods correspond to five-year intervals, and observations are weighted by total expenditure in the corresponding drug category in the MEPS.

Column 1 of Panel A shows that the NLLS estimate of α for all new drugs is 6.73 with a standard error of 2.33, which is significant at the 1 percent level.

The theoretical analysis suggests the possibility of entry of new drugs responding to anticipated changes in market size. The remaining columns of Panel A investigate whether it is current market size or past or future market sizes that have the strongest effect on entry of new drugs by including lags ($\log M_{c,t-1}$) and leads ($\log M_{c,t+1}$) of potential market size on the right hand side of the estimating equation.

Column 2 includes current and lagged market size together,[18] column 3 looks at the effect of lagged market size alone, column 4 includes both current and lead market size, and finally, column 5 looks at the relationship between lead market size and entry of new drugs. The entry of all drugs is most responsive to current or five-year lead market size. When current and lag market sizes are included together, current market size has a similar magnitude to column 1, while lag market size is negative, though insignificant, presumably because current and previous market sizes are highly correlated. When current and lead market sizes are included together, current market size is not significant, whereas lead market size is marginally significant. Moreover, column 5 shows that the five-year lead of market size has somewhat greater predictive power for entry of new drugs than current market size (the estimate of α is now 10.02, with standard error 2.61, and the R^2 of the regression has increased to 0.86 from 0.81 in column 1).

[18] We construct the lagged market size measures for 1960s using demographic information from the CPS, so the number of observations does not decline. The results are similar if we only use the post-1970 data.

The working paper version, Acemoglu and Linn [2003], reported results using a different classification of drugs based purely on age structure rather than drug indications. Appendix Table A2 contains estimates using this classification. The results are qualitatively similar, though quantitatively somewhat smaller. For example, the basic specification with all drug approvals in column 1 now leads to an estimate of α equal to 3.95 (standard error = 1.16), which is again significant at 1 percent.

5.2 Market Size and Non-Generics

The results in Panel A combine generics and non-generics. Entry of generics and non-generics may be driven by different processes. Moreover, generics, which are identical to existing drugs, do not correspond to "innovation". For the purposes of understanding the relationship between market size and innovation, the response of non-generics to potential market size is therefore more relevant.

Panel B shows a positive but statistically insignificant relationship between potential market size and entry of new non-generic drugs. However, column 5 shows that there is a strong and statistically significant relationship between five-year leads of market size and entry of new non-generic drugs. The estimate is 5.11, with standard error 2.22.[19] Appendix Table A2 shows that the results are similar with our alternative classification scheme; for example, the effect of five-year leads of market size on the entry of non-generic drugs is estimated to be 6.31 with standard error 2.18, which is significant at 1 percent. These results suggest that the entry of new non-generic drugs responds to anticipated future market size, perhaps with five-year leads, which is consistent with the possibility of limited anticipation effects highlighted by the theoretical model.[20]

Panel C shows an even stronger and larger effect of potential market size on the entry of generics. The estimate of α in the baseline specification of column 1 is 12.19 (standard error=3.29). The estimates in the other columns are similar, and suggest that the entry of generics respond most strongly to current market size. Appendix Table A2 shows that using the old classification system leads to similar results, though the magnitudes of the effect of market size on generic entry are much smaller. For example, the estimate in column 1 is 6.50

[19] The difference between the estimates in columns 1 and 5 reflects not the differences in samples but the response of the entry of non-generics to lead market size rather than current market size. Using the current market size with the same sample as the lead market size specification yields a coefficient of 2.54 (standard error = 1.72).

[20] Here "limited" does not refer to the strength of the effect, but to the fact that the response to market size is 5 years before the change in market size, not further in advance. The estimate with two-period (ten-year) lead of market size is 5.95 with standard error 3.11, while those with further leads are insignificant. Although the ten-year lead is marginally significant, we focus on the five-year lead since it is more precisely estimated and enables us to have a larger sample.

(standard error = 2.17).

The magnitude of the effects shown in Table II, in particular the impact of market size on non-generics, is large, but plausible. For example, the estimate in column 5 of Panel B implies that a 1 per cent increase in our market size measure leads to about a 5.1 percent increase in the entry of new non-generic drugs. There are a total of 2,203 non-generic approvals between 1970 and 2000, thus on average 10 approvals in every five-year interval in each of our 33 categories. Therefore, our estimate implies that a 2 percent increase in market size should lead to the entry of about 1 new drug. Total pharmaceutical sales were approximately $130 billion in 1999 [IMS, 2000], which implies an average annual expenditure of $3.9 billion per category. A 2 percent increase therefore corresponds to $78 million, or about $1.07 billion over 15 years, which is the life of a typical non-generic drug. Since entry costs for non-generics are around $800 million [in 2000 dollars, DiMasi et al, 2001], entry of one new drug in response to an increase of approximately $1 billion in revenue is within the range of plausible responses Naturally, this calculation is very rough and only suggestive, since it ignores the difference between average demand and the demand that a marginal entrant will capture.[21]

5.3 Robustness

Table III investigates the robustness of the results shown in Table II. Although our main focus is the response of new non-generic drugs to market size, we also show results for all approvals (Panel A) and for generics (Panel C). Since Table II shows that entry of new non-generic drugs responds to five-year leads of market size, Panel B looks at the robustness of the effect of lead market size on non-generic entry.

Column 1 replicates the baseline results from Table II (column 1 for Panels A and C, and column 5 for Panel B). In column 2, we use ten-year intervals instead of the five-year intervals. The estimate of α for all drug approvals is slightly smaller, 6.22 (standard error=2.21), while the estimate for non-generics is larger, 8.89, but less precise (standard error = 4.57). Column 3 looks at the effect of changes in market size driven purely by population changes, and shows very similar results to the baseline estimates (in this case, regression weights are given by total use in the corresponding category in the MEPS).

[21] It also has to be borne in mind that these estimates are informative about the effect of market size on the *composition* of research, and the relationship between total pharmaceutical market size and aggregate research could be quite different. If we estimate (12) for all approvals without time effects, we obtain a coefficient of 0.32, with a standard error of 0.11, on current market size (for non-generics, the estimate on lead market size is 0.40, with a standard error of 0.09). This is consistent with the view that the response of the composition of R&D to market size is quite different from the response of total R&D. Nevertheless, the difference between the results with and without time effects is at least partly due to the presence of other time-varying factors affecting entry of new drugs, for example, the fluctuations in FDA approvals, unrelated to market size, discussed in footnote 17.

Column 4 uses an OECD market size measure combining West European and Japanese demographic information with the U.S. information.[22] Since we only have information on population shares for the other countries, we perform this exercise for the population-based measure of market size. The U.S. and OECD populations by age group have a high correlation, equal to 0.81. Using the OECD market size measure leads to qualitatively similar, but quantitatively smaller results than those obtained using only the U.S. information. For example, for all approvals, the estimate of α is 4.24 (standard error=1.36), while for non-generics, the estimate is 3.43 (standard error = 1.83). The smaller estimate with the OECD market size measure might be because entry of new drugs in the United States is more responsive to the U.S. market size, or because there is greater measurement error in our OECD market size estimates, which are constructed from only population data, available at five-year intervals, combined with age profiles for U.S. consumers.

Column 5 investigates the effect of weighting on the estimates. The unweighted estimates of α for all approvals and for generics are smaller than the baseline. For example, for all approvals, the estimate is 3.32 (standard error = 1.74). But the unweighted estimate for non-generics is very similar to the baseline, 4.57 (standard error = 1.91), and continues to be significant at 5 percent.

Column 6 uses an alternative measure of market size constructed with single age groups for i_{at}'s and u_{ca}'s in equation (13). This procedure uses more information about the age profiles, but since there are fewer observations in some single age groups, the estimates of u_{ca}'s are less precise. The estimates using this alternative measure are very similar to the baseline results. For example, the estimates of α for non-generics is 5.01 with a standard error of 2.20.

Column 7 estimates the model in (10) with the Pakes-Griliches transformation using OLS. The estimates are similar to those in column 1. For example, for non-generics, the estimate of α is 4.53, with standard error 2.24, and is significant at 5 percent.

Finally, columns 8, 9 and 10 check the robustness of the results to dropping some of the categories that are most heterogeneous in terms of types of drugs and expenditure profiles. These are the Antibiotic, Cardiac, and Pain Relief categories. In all cases, the exclusion of these categories has little effect on the estimates.[23]

[22] Obtained from the United Nations website, esa.un.org/unpp/.

[23] We have also experimented with dropping each of the other categories. The effect of market size on all approvals remains significant at 5 percent in all cases. With two exceptions, Acid/Peptic Disorders and Estrogens/Progestins, dropping any of the other categories also makes little difference for non-generics. When these categories are dropped, the effect of lead market size on non-generic entry is somewhat smaller and only significant at 10 percent.

5.4 Potential Supply-Side Determinants of Innovation

The first part of Table IV investigates the robustness of the baseline results to controlling for potential non-profit determinants of innovation, such as changes in scientific incentives or opportunities captured by the δ_j's in the theoretical model. Although our focus is on the effect of lead market size on the entry of new non-generic drugs (Panel B), for completeness, we also report the estimates of the effect of current market size on all approvals and on generics separately (Panels A and C).

First, recall that the major threat to our identification strategy is changes in the δ_j's (since permanent differences in δ_j's are already taken out by the drug category fixed effects). If the δ_j's change over time, they are also likely to be serially correlated. Adding lags of $\log N_{ct}$ to our basic specifications is therefore a simple way to check the importance of these concerns.

Column 1 of Table IV reports the results of estimating a lagged dependent variable specification, by adding a one-period lag of the dependent variable, $\log N_{ct-1}$, to our basic specification. The basic estimating equation changes from (12) to:

$$S_{ct} = \frac{\exp\left(\alpha \cdot \log M_{ct} + \psi \cdot \log N_{ct-1} + \mu_t\right)}{\sum_{\tau=1}^{T} \exp(\alpha \cdot \log M_{c\tau} + \psi \cdot \log N_{c\tau-1} + X'_{c\tau} \cdot \beta + \mu_\tau)} + \varepsilon_{ct}, \tag{14}$$

where recall that $S_{ct} = N_{ct} / \sum_{\tau=1}^{T} N_{c\tau}$. Since $\log N_{ct-1}$ is correlated with the error term, estimates of this equation would be biased. To deal with this problem, we instrument $\log N_{ct-1}$ with $\Delta \log N_{ct-2}$. This is a valid instrument as long as there is no additional autocorrelation in the error term, ε_{ct}.[24] This specification is also useful to check for other sources of serial correlation in the entry rate of new drugs.

The estimate of α from equation (14) for non-generics, reported in column 1 of Panel B, is similar to the estimate in Table II, 5.05 with a standard error of 2.19, and the lagged dependent variable is insignificant (0.01, with standard error 0.23). Similarly, if we look directly at the residuals from the estimation of (12), there is no evidence of residual serial correlation for non-generic drugs. These results therefore show no evidence of significant residual serial correlation due to changes in scientific opportunities or other reasons, and also demonstrate that controlling for serial correlation has little effect on our estimates for non-generics.

The situation is different for generics. In this case, the lagged dependent variable is highly significant, and the estimate of α is much smaller, imprecise and insignificant.[25] When we

[24] See, for example, Arellano and Bover [1995] and Blundell and Bond [1998]. We cannot use other commonly-used moment restrictions, since equation (11) cannot be first-differenced. As with the specifications with lag market size in Table II, we use information on approvals before 1970 to construct lags, so the sample size remains the same as in the basic specification.

[25] Since generic drugs can only enter after patents on non-generics expire, a more satisfactory specification

combine generics and non-generics and look at all approvals in Panel A, the lagged dependent variable is still significant, and the estimate of α is smaller, though still significant at 5 percent. These results show that there is significant residual serial correlation for generics, presumably because a process of slow entry of new generics after patents expire [see Reiffen and Ward, 2004]. This reduces our confidence in the magnitude of the estimates for generics and all approvals. The rest of the table continues to show the other robustness results for generics and all approvals for completeness, though we do not dwell on these.

A plausible conjecture is that non-profit incentives to develop drugs would be particularly responsive to opportunities to save lives or cure major illnesses. Our second strategy looks at variation in the health benefits of new drugs across categories. New drugs in our data set include both drugs that are demanded by the consumers but do not "save lives", such as Prozac, Paxil, Vioxx, or Viagra, and those that actually save lives such as heart medicines or cancer treatments (see Lichtenberg [2002, 2003], on the effect of pharmaceutical innovations on declines in mortality). To investigate this issue, we measure the number of life-years lost corresponding to each drug category using the Mortality Detail Files from the National Center for Health Statistics from 1970-1998. Following Lichtenberg [2002], for each death, we subtract the person's age from 65, then calculate the total number of life-years lost for all the deaths resulting from diseases related to drugs in each category.[26] Column 2 adds this measure of life-years lost to the right hand side of our baseline regression models as a proxy for this source of non-profit incentive to undertake research. The estimate of the effect of lead market size on non-generics is now 4.85 (standard error = 2.33).

Third, we investigate the implications of differences in scientific funding for various drug categories. Using the Computer Retrieval of Information on Scientific Projects (CRISP) dataset (details are contained in Lichtenberg [2001] and Acemoglu and Linn [2003]), we construct a variable measuring the total amount of federal funding for research projects in all drug categories, and include this variable as a control on the right hand side. To the extent that government funding also responds to potential market size (for example, because drug companies have a greater tendency to apply for funding in areas where they plan to do research), this variable would be correlated with our market size measure. In practice, the correlation is low, and col-

would include both lagged generic and non-generic approvals on the right hand side. In this case, when estimated without instrumenting, the effect of market size on entry of generics has a reasonable magnitude 6.31 (s.e.=4.46). The coefficient on lagged non-generic approvals is 0.34 (s.e.=0.18) and that on lagged generic approvals is 0.39 (s.e.=0.15). This specification is not instrumented, since we do not have generic approvals before 1970, and instrumentation would reduce the number of time periods substantially. Since generics are not our main focus, we do not pursue this specification further.

[26] For example, if someone dies at age 32, this counts as 33 life years lost; people dying older than 65 receive no weight in this calculation. Since many of our categories contain diseases or conditions that do not lead to death, we obtain several empty cells.

umn 3 shows that the inclusion of this variable has little effect on our estimates. The estimate of the effect of lead market size on new non-generics is 5.68 (standard error = 2.25), though the funding variable itself is also significant at 5 percent (not reported in the table), and shows that the availability of funding might also have an effect on innovation.

Fourth, to control for potential trends in scientific opportunities across drug categories, we add proxies for pre-existing trends. We construct an estimate for pre-existing trends as $\Delta_c = (\log N_{c,70} - \log N_{c,40})/6$, where $\log N_{c,70}$ is the log approvals for category c in 1970 and $\log N_{c,40}$ is the log approvals in 1940. We then estimate the equation:

$$S_{ct} = \frac{\exp\left(\alpha \cdot \log M_{ct} + \Delta_c \cdot \sigma_t + \mu_t\right)}{\sum_{\tau=1}^{T} \exp(\alpha \cdot \log M_{c\tau} + \Delta_c \cdot \sigma_\tau + X'_{c\tau} \cdot \beta + \mu_\tau)} + \varepsilon_{ct}, \tag{15}$$

where σ_t's are such that $\sigma_t = 0$ if $t = 1970$, $\sigma_t = \sigma_{75}$ if $t = 1975$, $\sigma_t = \sigma_{80}$ if $t = 1980$, and so on. This specification allows drug categories that have grown at different rates between 1940 and 1970 to also grow at different rates in the later periods. Column 4 reports the results of this exercise. The estimates of α are similar to our baseline estimates; for non-generics, it is 5.59 (standard error = 2.38). These results are perhaps not surprising, since pre-1970 approvals are considerably noisier, thus only an imperfect control for pre-existing trends. We do not report this specification for generics, since there is no entry of generics prior to 1970.

An alternative, and substantially more demanding, strategy is to include in-sample linear time trends. To do so, we estimate:

$$S_{ct} = \frac{\exp\left(\alpha \cdot \log M_{ct} + \eta_C \cdot t + \mu_t\right)}{\sum_{\tau=1}^{T} \exp(\alpha \cdot \log M_{c\tau} + \eta_C \cdot t + X'_{c\tau} \cdot \beta + \mu_\tau)} + \varepsilon_{ct}, \tag{16}$$

where c refers to the 33 detailed drug categories, and C refers to the relevant 16 major drug categories, i.e., the one to which the detailed category c belongs to. We expect technological differences to be well approximated by the 16 major drug categories, which are based on broad therapeutic intent. The estimates, reported in column 5, are in fact larger than our baseline estimates. For example, for non-generics the estimate of α is 9.37 (standard error =4.22), and the linear trends are jointly significant at 1 percent. These results indicate that market size is unlikely to be proxying for differential scientific trends across the major drug categories.

We also investigate the potential effects of advances in biotechnology, such as the use of recombinant DNA, or other technological changes, during the late 1980s and the 1990s. In terms of our model, these developments would correspond to changes in the δ_j's. In column 6, we drop the categories of Cancer and Vascular, which, according to the FDA approval list, have witnessed the entry of the greatest number of orphan drugs, presumably by biotechnology firms. Even though, as noted above, our dependent variable does not include these drugs,

we also check whether our results are driven by entry of new drugs in these categories. The estimates in column 6 are close to those in column 1 of Table II.

In addition, there is anecdotal evidence that biotechnology firms were first active in producing insulin (the Glucose and Thyroid category) and in the Anemia category.[27] In column 7, we drop these two categories, and again find that our results are essentially unchanged. To assess the role of biotechnology firms further, we add the approvals of a group of products known as biologics, where biotechnology firms have been particularly active, to our measure of drug approvals. These products, which include some vaccines, blood and plasma related products, such as interferon and errythropoietins (used for red blood cell production), are not included in our baseline measure because they go through a separate FDA regulatory process. The results of this regression, reported in column 8, show little change in the estimates of α.

Finally, to see whether the advent of biotechnology or other technological advances of the past two decades have changed the relationship between market size and entry of new drugs, we estimate our baseline models including an interaction between a post-1985 (or post-1990) dummy and market size. Our estimates show no evidence of significant interactions. For example, in a specification parallel to the model for non-generics in column 5 of Table II (not reported), the estimate of α is 6.09 (s.e.=2.38), and the interaction with the post-1985 dummy is -0.13 (s.e.=0.09), thus small and insignificant.

The results in this subsection therefore show that a number of controls for non-profit factors have little effect on our main finding regarding the effect of lead market size on entry of new non-generic drugs. Although these results are not conclusive on the effect of scientific or other non-profit considerations in pharmaceutical research, they suggest that the effect of potential market size on entry and innovation is relatively robust.

5.5 Changes in Health Insurance Coverage

Our market size measure only exploits changes in potential market size driven by demographic trends. Another source of variation in market size comes from changes in coverage of drug expenditure in private or public health insurance programs. During our sample period, there were significant changes in the coverage of drug expenditure in health insurance plans. For example, the percentage of 60+ year-olds with private insurance rose from 60 percent to 75 percent between 1974 and 1996 (authors' calculations). We now investigate whether we can improve our measure of potential market size by including information on health insurance coverage.

[27] Biotechnology firms were also active in producing human growth factor, but since there are only a small number of individuals using these drugs in the MEPS, these drugs are not included in our approvals dataset.

We use the National Health Interview Survey (NHIS, 1974-1996) to construct a market size measure incorporating information on health insurance coverage as follows: $\widetilde{M}_{ct} = \sum_a u_{ca} \cdot i_{at} \cdot f_{at}$, where f_{at} is the fraction of those of age a in period t with private health insurance, u_{ca} and i_{at} are as defined above. Because there is no consistent information on prescription drug coverage, we assign prescription coverage to any individual with both doctor and surgical coverage. Prescription drug coverage is highly correlated with this measure in the years we can observe it. In column 9, we use $\log \widetilde{M}_{ct}$ as our market size measure instead of $\log M_{ct}$. This leads to similar, and somewhat more precise, results. For non-generics, the estimate of α is 4.45 with standard error 1.38. Despite the greater precision of these results, we have more confidence in our baseline estimates, since the measure $\widetilde{M}_{ct}$ effectively assigns 0 expenditure to those without insurance and relies on information on drug coverage imputed from doctor and surgical coverage.

5.6 Reverse Causality

Lichtenberg [2002, 2003] shows that new drugs have increased the average age at death. This introduces the potential for reverse causality whereby the market size for successful drugs may be endogenously larger, because their users live longer. This is unlikely to be a first-order concern, since drug-induced changes in population are small relative to the demographic changes that we are exploiting. Nevertheless, we further address this issue by instrumenting for current population using the corresponding population from 5 years before. For example, we use the income of 50-54 year-olds in 1975 as an instrument for the income of 55-59 year-olds in 1980. The fraction of 50-54 year-olds is highly correlated with the fraction of 55-59 year-olds 5 years later, but is unaffected by new drugs that are developed in the intervening 5 years.

These instrumental-variables estimates, reported in column 10 of Table IV, are similar to the baseline results and show no evidence of reverse causality. For example, for non-generics, the estimate in Panel B is 5.32 (standard error = 2.30).

5.7 New Molecules and Patents

The results so far show a large and robust effect of potential market size on entry of new non-generic drugs, and suggest a strong link between market size and innovation. In this subsection, we briefly investigate the relationship between market size and some other measures of pharmaceutical innovation, namely, the introduction of new molecules and patents.[28]

[28] We were unable to obtain data for a sufficient number of categories for another possible proxy for pharmaceutical innovation, clinical trials.

First, the FDA has classified some new drugs as containing a new molecule (an active ingredient that has not been marketed before in the United States). In all, there are 442 new molecules in our dataset compared to 2,203 non-generics, indicating that new molecules may correspond to significantly more important innovations. We use the introduction of new molecules as an alternative measure of innovation. Panel D of Appendix Table A2 shows that there is a positive but insignificant association between current market size and new molecules, and a positive and statistically significant relationship between lead market size and new molecules. The estimate of α with lead market size is 4.35 (standard error = 2.18) and is significant at 5 percent.

Finally, we investigate the effect of potential market size on patents. We obtained data on pharmaceutical patents from Thomson Derwent Inc., and with the help of a specialist at this company, we mapped these patents into our classification system.[29] However, the mapping of patents for chemicals into the FDA and disease categories is imperfect and necessarily introduces a significant amount of noise, which makes inference more difficult in this case.

Firms typically apply for a patent prior to the clinical trial stage of drug development, or about 5-10 years before the drug is approved.[30] Given the results so far, we might expect patents to respond to future demographic changes. Panel D of Appendix Table A2 shows that there is no relationship between current or lead market size and new patents.

There may be a number of reasons for the weaker results with patents.[31] First, this may simply reflect the imperfect match between the patent data and the FDA categories, especially bearing in mind the potential use of certain chemical structures in multiple drug lines. Second, the significant costs and uncertainty involved in the development of new molecules and patentable products may be creating substantial attenuation (e.g., a drug intended for the 1990s may be patented in the 1980s or 1990s, depending on delays in the research process). Third, pharmaceutical companies may respond more to profit incentives at the later stages of the research process than at the earlier stages. Finally, patents may be more responsive to OECD demand than to U.S. demand. To investigate the last possibility, column 5 of Panel D of Appendix Table A2 looks at the relationship between changes in OECD market size derived from European, Japanese and U.S. demographic changes. In this case, we find a significant

[29] We could not use the data from the Hall-Jaffe-Trachtenberg patent dataset (see Jaffe and Trachtenberg, 2002) because we were unable to map their classification based on chemical structure to our drug categories.

[30] The firm therefore loses a significant fraction of the life of the patent before it can begin marketing the drug. Part of the 1984 Hatch-Waxman Act allowed pharmaceutical companies to apply to the FDA for an extension of the life of their patents, if they could show that they lost marketing time while waiting for approval. The maximum extension is 5 years, and depends, among other things, on the length of the initial FDA approval process. Overall, companies have a maximum of 14 years of patent protection after FDA approval.

[31] Finkelstein [2003] also finds weaker results for vaccine patents than for later stages of development.

relationship between market size and patents. With lead OECD market size, the estimate of α is 4.60 (standard error = 1.85).[32] Although this result suggests that OECD demand may be more important for patents, we are currently unable to make more progress in distinguishing between these various explanations, and the weaker results for patents remain a puzzle.

6 Concluding Remarks

This paper investigates the response of entry of new drugs and pharmaceutical innovation to changes in potential market size of users, driven by demographic changes. Our results indicate that a 1 percent increase in the potential market size for a drug category leads to approximately a 5 percent growth in the entry of new non-generic drugs approved by the FDA. Entry of non-generic drugs appears to respond to increases in future market size in the next five years or so. There is also a substantial response of entry by generic drugs mostly to current market size.

The relationship between market size and entry of new drugs, if further proven to be robust, has important implications both for research on the pharmaceutical industry, and for the endogenous growth and directed technical change literatures. It provides evidence that, as conjectured by these models, R&D and technological change are directed towards more profitable areas. Furthermore, the magnitude of the effect, which is important for evaluating various theoretical predictions of these models, is substantial. For example, directed technical change models suggest that the relative demand curves for factors of production can be upward, rather than downward, sloping if the development of new technologies responds to a 1 percent increase in market size by more than 1 percent (see, for example, equations (21) and (22) in Acemoglu, [2002])—the corresponding number implied by our estimates is around 5. Second, these findings imply that pharmaceutical research towards drugs with relatively small markets may be limited, which is a key premise of recent work by Kremer [2002]. Kremer suggests that there needs to be selective government incentives for developing drugs against malaria and other third-world diseases.

We view this research as part of a broader investigation of the effects of profit incentives on innovation. Evidence from a single industry may be nonrepresentative, for example because the pharmaceuticals may be more research oriented than other industries. Future research investigating the response of innovation and entry of new products to market size both in specific industries and at the economy-wide level is necessary to substantiate the results presented here.

[32] Part of the difference between the U.S. and OECD results is driven by the fact that we are using income-based measures for the U.S. and the population-based measures for the OECD. Using the population-based measure for the U.S. lead market size yields a coefficient of 2.79 (standard error = 1.97).

7 References

Acemoglu, Daron. 1998. "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality." *Quarterly Journal of Economics*. 113, pp. 1055-1090.

Acemoglu, Daron. 2002. "Directed Technical Change" *Review of Economic Studies.* 69, pp. 781-810.

Acemoglu, Daron. 2003. "Labor- and Capital-Augmenting Technical Change." *Journal of European Economic Associations*, 1, pp. 1-37.

Acemoglu, Daron and Joshua Linn. 2003. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry" NBER Working Paper No. 10038.

Aghion, Philippe and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction" *Econometrica*, 60, pp. 323-351.

Arellano, Manuel and Olympia Bover. 1995. "Another Look at the Instrumental-Variable Estimation of Error-Components Models" *Journal of Econometrics*, 68, pp. 29-52.

Arrow, Kenneth J. 1962. "The Economic Implications of Learning-by-Doing," *Review of Economic Studies*, 29, pp. 155-173.

Blundell, Richard and Stephen Bond. 1998. "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models" *Journal of Econometrics*, 87, pp. 115-144.

Cerda, Rodrigo. 2003. "Drug, Population and Market Size" University of Chicago Ph.D. dissertation.

Cockburn, Rebecca M. and Iain M. Henderson. 2001. "Scale and Scope in Drug Development: Unpacking the Advantages of Size in Pharmaceutical Research." *Journal of Health Economics*, 20, pp. 1033-1057.

Danzon, Patricia M., Sean Nicholson and Nuno Sousa Pereira. 2003. "Productivity in Pharmaceutical-Biotechnology R&D: The Role of Experience in Alliances." *NBER Working Paper* #9615.

DiMasi, Joseph A., Ronald W. Hansen and Louis Lasagna. 1991. "Cost of Innovation in the Pharmaceutical Industry." *Journal of Health Economics*, 10, pp. 107-142.

DiMasi, Joseph A., Ronald W. Hansen and Henry G. Grabowski. 2003. "The Price of Innovation: New Estimates of Drug Development Costs." *Journal of Health Economics*, 22, pp. 151-185.

Drandakis, E. and Edmund Phelps. 1965. "A Model of Induced Invention, Growth and Distribution" *Economic Journal*, 76, pp. 823-840.

Dranove, David and David Meltzer. 1994. "Do Important Drug Reach the Market Sooner?" *RAND Journal of Economics*, 25, pp. 402-423.

Finkelstein, Amy. 2003. "Health Policy and Technical Change: Evidence from the Vaccine Industry." NBER Working Paper No. 9460.

Galambos, Louis and Jeffrey Sturchio. 1998. "Pharmaceutical Firms and the Transition to Biotechnology: A Study in Strategic Innovation." *Business History Review*, 72, pp. 250-278.

Gambardella, Alfonso. 2000. *Science and Innovation: The U.S. Pharmaceutical Industry During the 1980s*, Cambridge University Press, Cambridge UK.

Griliches, Zvi. 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica*, 25, pp. 501-522.

Grossman, Gene and Elhanan Helpman. 1991. *Innovation and Growth in the Global Economy.* MA: MIT Press.

Hausman, Jerry, Bronwyn H. Hall and Zvi Griliches. 1984. "Econometric Models for Count Data with an Application ot the Patents-R&D Relationship." *Econometrica*, 52, pp. 909-938.

Hayami, Yujiro and Vernon Ruttan. 1970. "Factor Prices and Technical Change in Agricultural Development: the U.S. and Japan, 1880-1960" *Journal of Political Economy*, 77. pp. 1115-1141.

Henderson, Rebecca and Iain M. Cockburn. 1996. "Scale, Scope and Spillovers: The Determinants of Research Productivity in Drug Discovery." *Rand Journal of Economics,* 27, pp. 32-59.

IMS. 2000. "World Review" at www.ims-global.com/insight/report/market_growth/report0600.htm.

Jaffe, Adam and Manuel Trachtenberg and Rebecca Henderson. 2002. *Patents, Citations, and Innovations*. MA: MIT Press..

Kennedy, Charles 1964. "Induced Bias in Innovation and the Theory of Distribution" *Economic Journal,* 74, pp. 541-547.

Kling, Jim. 1998. "From Hypertension to Angina to Viagra." *Modern Drug Discovery,* 1 (2), pp. 31-8.

Kremer, Michael. 2002. "Pharmaceuticals and the Developing World" *Journal of Economic Perspectives*. 16 (4), pp. 67-90.

Lichtenberg, Frank R. 2001. "The Allocation of Publicly Funded Biomedical Research," in *Medical Care Output and Productivity, Studies in Income and Wealth* volume 63. Ernst Berndt and David Cutler, eds. University of Chicago Press, 2001.

Lichtenberg, Frank R. 2002. "Sources of U.S. Longevity Increases, 1960-1997." *NBER Working Paper* #8755.

Lichtenberg, Frank R. 2003. "The Impact of New Drug Launches on Longevity: Evidence from Longitudinal, Disease-level Data from 52 Countries, 1982-2001". Columbia mimeo.

Lichtenberg, Frank R. and Suchin Virabhak 2002. "Pharmaceutical Embodied Technical Progress, Longevity and Quality of Life: Drugs as Equipment for Your Health." *NBER Working Paper* #9351.

Lichtenberg, Frank R. and Joel Waldfogel. 2003. "Does Misery Love Company? Evidence From Pharmaceutical Markets Before and After the Orphan Drug Act." #9351

Ling, Davina C., Ernie Berndt, and Richard G. Frank. 2003. "General Purpose Technologies, Capital-Skill Complementarity, and the Diffusion of New Psychotropic Medications among Medicaid Populations" MIT mimeo.

Malerba, Franco and Luigi Orsenigo. 2002. "Innovation and Market Structure in the Dynamics of the Pharmacetutical Industry: Towards a History Friendly Model." *Industrial and Corporate Change*. 11 (4), pp. 667-703.

Newell, Richard, Adam Jaffee and Robert Stavins. 1999. "The Induce Innovation Hypothesis and Energy-Saving Technological Change." *Quarterly Journal of Economics*. 114, pp. 907-940.

Pakes, Ariel and Zvi Griliches. 1980. "Patents and R&D at the Firm Level: A First Look." *Economic Letters*, 5, pp. 377-381.

Pakes, Ariel and Mark Schankerman. 1984. "An Explanation into the Determinants of Research Intensity." pp. 209-232 in *R&D, Patents and Productivity* edited by Zvi Griliches, National Bureau of Economic Research, University of Chicago Press, Chicago 1984.

Popp, David. 2002 "Induced Innovation and Energy Prices" *American Economic Review*. 92, pp. 160-180.

Reiffen, David and Michael R.Ward. 2004 "Generic Drug Industry Dynamics." forthcoming *Review of Economics and Statistics*.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy*, 98. S71-S102.

Samuelson, Paul. 1965. "A Theory of Induced Innovations Along Kennedy-Weisacker Lines" *Review of Economics and Statistics*, 47, pp. 444-464.

Schmookler, Jacob. 1966. *Invention and Economic Growth*. MA: Harvard University Press.

Scott Morton, Fiona. 1999. "Entry Decisions in the Generic Drug Industry." *The Rand Journal*, 30, pp. 421-440.

Wooldridge, Jeffrey M. 1999. "Distribution-free Estimation of Some Nonlinear Panel Data Models." *Journal of Econometrics*, 90, pp. 77-97.

Wooldridge, Jeffrey M. 2002. *Econometrica Analysis of Cross Section and Panel Data*, MIT Press Cambridge.

8 Data Appendix

8.1 Medical Expenditure Panel Survey (MEPS) and Construction of the Drug Classification System

The MEPS is an annual survey of randomly sampled households; we use the 1996, 1997 and 1998 surveys. We obtain each person's age, the name and national drug code of the prescription drug(s) used, and total expenditure (there are multiple records for people who use more than one prescription drug). Over the 3 years, we have about 500,000 drugs used and about 85,000 people. Expenditure information includes out-of-pocket expenses, as well as amounts paid by insurance companies and government payments (e.g., Medicaid and worker's compensation). These information are collected from the pharmacies and medical providers listed by the respondents.

We begin with the 159 therapeutic categories, obtained from the FDA's National Drug Code (NDC) Directory. The names of these categories can be found in the second column of Appendix Table A1. The NDC Directory contains a file with the therapeutic category for most FDA approved drugs currently on the market. We assign each drug in the MEPS to one of the 159 categories by matching it by national drug code with a drug in the NDC file. We cannot match about 10 percent of the drugs mentioned in the MEPS; these are usually not commonly used drugs, and make up less than 5 percent of the total drugs used.

Drug expenditure shares and use per person are calculated by computing drug expenditure and use by five-year age group, and then dividing by the income and population estimates for the same age group from the CPS. The results are also very similar if we construct expenditure shares (use per person) as a weighted average of expenditure shares (use per person) for individuals in the survey (i.e., without using CPS information). For example, although the estimates of the effect of current market size on all approvals are somewhat smaller, the estimate of the effect of lead market size on the entry of non-generics, which is our main focus, is similar to our baseline estimates, 4.88, with standard error 1.54. We prefer to use income estimates from the CPS because the MEPS income data are likely to contain greater measurement error; in the MEPS, wage and salary incomes for almost half of the sample are imputed either based on broad income ranges or other information, and non-wage incomes were generally imputed and also the imputation methods changed between the 1996-97 and 1998 surveys. We also checked the robustness of our results using an alternative market size measure constructed with single age groups, and the results are reported in Table III. We prefer the measure using five-year age groups, since there are only a few observations in some single-age groups in the MEPS.

The FDA has assigned the 159 categories to one of 20 major therapeutic categories. As noted in the text, we drop four major categories: Anesthetics, Antidotes, Radiopharmaceuticals, and Miscellaneous.[33] Within each major category, we first separate categories whose drugs

[33] We also drop several minor categories when there are not sufficient observations to estimate a reliable age structure. We use about 1,000 observations as our cutoff rule. We obtain this number from observing that only

have different indications (we determine drugs' indications by searching by name on the National Institute of Health website, www.nlm.nih.gov/medlineplus/druginformation.html). For categories that have not been separated based on indications, we then separate them if there is sufficient heterogeneity in the age profile of users for subcategories. The categories in Table A1 make it clear that we create subcategories when there is considerable age variation within broad categories. For example, Antimicrobials are separated into category 1, Antibiotics, used mostly by 0-20 year olds; category 2, Antivirals, used predominantly by individuals of age 30-60; and category 4, Antifungals, which have a balanced age profile. This classification system differs somewhat from the working paper version, in which we divided major classes based entirely on age structure. The details of the previous classification system are in Acemoglu and Linn [2003], and Appendix Table A2 reports results using this older classification.

8.2 Drug Approvals from the FDA

The list of FDA drug approvals were obtained by Lichtenberg and Virahbak [2002] under the Freedom of Information Act. We thank Frank Lichtenberg for generously sharing these data with us. Over-the-counter drugs and orphan drugs (of which only a few can be matched) are excluded. Biologics, which go through a separate approval process, are not in this dataset.

We match drugs in the approval list to FDA categories by drug name and FDA approval number. 14,432 of 16,772 prescription drugs (86 percent) approved since 1970 are matched, while before 1970, the match rate is about 51 percent. This motivates our focus on drug approvals between 1970 and 2000. Drugs that have the same approval number as a previously approved drug and drugs for which the corresponding FDA category is dropped because of insufficient observations in the MEPS are excluded. Finally, we drop drugs with the same name, MEPS category and company has a previously approved drug, leaving us with our sample of 5,374 drugs. Of these, 2,203 are non-generics, and 442 new molecules.

8.3 Patents

We have obtained patent data from Thomson Derwent Inc. We use all pharmaceutical patents granted in the United States, between 1970-2000. We use these data instead of the Hall-Jaffe-Trachtenberg patent data because the latter use a classification for pharmaceuticals based on chemical structure, which we are not able to map into our FDA classification system. The Thomson Derwent patents are classified by chemical structure and therapeutic intent, and a specialist at the company has mapped this system into the FDA system. The mapping is not precisely one to one, and we drop about five percent of the patents that fall into two of our categories. We are left with 275,406 patents.

categories with more than 1,000 observations have fairly smooth age profiles.

TABLE I

Correlations Between Different Drug Use Measures

	Panel A: NAMCS over time		
	1980/1990	1990/2000	1980/2000
Correlation	0.897	0.861	0.861
Weighted Correlation	0.906	0.843	0.856
Mean Correlation by Drug	0.709	0.651	0.626
Weighted Mean Correlation by Drug	0.820	0.825	0.790
	Panel B: MEPS over time		
	1996/1997	1997/1998	1996/1998
Correlation	0.961	0.965	0.929
Weighted Correlation	0.962	0.973	0.937
Mean Correlation by Drug	0.698	0.686	0.575
Weighted Mean Correlation by Drug	0.865	0.881	0.796

	Panel C: NAMCS/MEPS and MEPS use/expenditure	
	MEPS/NAMCS	MEPS use/MEPS expenditure
Correlation	0.869	0.954
Weighted Correlation	0.891	0.956
Mean Correlation by Drug	0.804	0.902
Weighted Mean Correlation by Drug	0.935	0.940

Notes: Correlation refers to the correlation of use per person or average expenditure share between the indicated dates and datasets. In weighted correlations, observations are weighted by total use or expenditure from the MEPS or NAMCS. Mean correlation by drug computes correlations separately by drug, then calculates the average.

TABLE II

Effect of Changes in Market Size on New Drug Approvals

	(1)	(2)	(3)	(4)	(5)
	Panel A: NLLS for Poisson model, dep var is count of drug approvals				
Market Size	6.73 (2.33)	8.76 (5.58)		1.86 (4.00)	
Lag Market Size		-1.83 (4.32)	4.94 (1.81)		
Lead Market Size				8.07 (5.13)	10.02 (2.65)
R Squared	0.81	0.81	0.80	0.86	0.86
	Panel B: NLLS for Poisson model, dep var is count of non-generic drug approvals				
Market Size	2.45 (2.19)	8.95 (5.11)		3.00 (4.05)	
Lag Market Size		-5.87 (4.53)	0.92 (2.07)		
Lead Market Size				1.88 (4.94)	5.11 (2.22)
R Squared	0.83	0.83	0.83	0.84	0.84
	Panel C: NLLS for Poisson model, dep var is count of generic drug approvals				
Log Market Size	12.19 (3.29)	6.16 (8.22)		13.87 (6.82)	
Lag Market Size		5.52 (6.24)	10.30 (2.32)		
Lead Market Size				-2.08 (8.43)	12.49 (4.53)
R Squared	0.70	0.70	0.70	0.70	0.69
Number of Observations	198	198	198	165	165

Notes: Huber-White robust standard errors are reported in parentheses. The dependent variable in Panel A is count of drug approvals, in Panel B the dependent variable is count of non-generic drug approvals, and in Panel C, it is generic drug approvals, all calculated from the FDA dataset of New Drug Approvals (see Appendix). Market Size is log potential market size calculated from the MEPS and the CPS (see text). Lag Market Size refers to one-period lag of Market Size, and Lead Market Size refers to one-period lead of Market Size. All regressions include drug and time dummies, an use the income-based measure of market size. Time intervals are 5 years. Estimates are weighted by total expenditure for the category in the MEPS. The Poisson model is estimated by Non-Linear Least Squares (with the Hausman, Hall and Griliches [1984] transformation). See equation (12) in the text.

TABLE III

Robustness Checks

	baseline NLLS	10 year intervals	population-based market size	OECD market size	unweighted regressions	market size uses single age groups	OLS	drop Antibiotics	drop Cardiac	drop Pain Relief
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: dependent variable is count of drug approvals									
Market Size	6.73	6.22	6.33	4.24	3.32	6.11	5.44	6.77	7.14	6.72
	(2.33)	(2.21)	(2.46)	(1.36)	(1.74)	(2.25)	(1.94)	(2.67)	(2.66)	(2.43)
R Squared	0.81	0.89	0.84	0.84	0.72	0.80	0.83	0.80	0.79	0.79
Approvals	5374	5374	5374	5374	5374	5374	5374	4651	5019	4815
	Panel B: dependent variable is count of non-generic drug approvals									
Lead Market Size	5.11	8.89	6.65	3.43	4.57	5.01	4.53	6.25	4.90	5.27
	(2.22)	(4.57)	(2.82)	(1.83)	(1.91)	(2.20)	(2.24)	(2.25)	(2.48)	(2.32)
R Squared	0.84	0.89	0.85	0.85	0.73	0.84	0.88	0.84	0.83	0.83
Approvals	1745	1745	1745	1745	1745	1745	1745	1174	1646	1648
	Panel C: dependent variable is count of generic drug approvals									
Log Market Size	12.19	11.68	14.36	8.95	5.96	11.21	7.85	11.74	13.25	12.30
	(3.29)	(2.91)	(3.27)	(2.00)	(2.81)	(3.10)	(2.78)	(3.79)	(3.75)	(3.47)
R Squared	0.70	0.79	0.75	0.75	0.54	0.70	0.79	0.69	0.69	0.69
Approvals	3171	3171	3171	3171	3171	3171	3171	3085	2941	2759

otes: Huber-White standard errors in parentheses. Dependent variables are count of drug approvals in Panel A, count of non-generic approvals in Panel B, and count of generic approvals in Panel C, all computed from the FDA dataset of New Drug Approvals. Panels A and C use current market size, and Panel B uses one period lead market size. All columns except 3 and 4 use the income-based measure of market size; columns 3 and 4 use the population-based measure. All regressions include drug and time dummies, and, except for column column 5, are weighted. Market size in column 4 is computed using total OECD population, as explained in text. In column 6 market size is computed using single year age groups (see text). The Poisson model is estimated using Non-Linear Least Squares (with the Hausman, Hall and Griliches, 1984, transformation) in columns 1-6 and 8-10.In column 7 if a cell is empty, the log approvals variable is set equal to zero, and a dummy variable equal to 1 when the cell is empty is added (see text).

TABLE IV

Potential Supply-Side and Demand-Side Determinants of Innovation

	lag dep var	life years lost	CRISP	pre-existing trends	major cat trends	drop Cancer, Vascular	drop Thyroid, Anemia	include Biologics	health insurance	IV with prev mkt size
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: NLLS for Poisson model, dependent variable is count of non-generic drug approvals									
Market Size	3.98	6.01	6.73	8.05	9.02	8.03	6.88	6.71	3.46	6.73
	(2.02)	(2.45)	(2.30)	(2.31)	(4.15)	(2.50)	(2.39)	(2.36)	(1.30)	(2.34)
Lagged Dependent Variable	0.97									
	(0.33)									
R Squared	0.86	0.81	0.81	0.85	0.91	0.79	0.80	0.81	0.80	0.81
	Panel B: NLLS for Poisson model, dependent variable is count of non-generic drug approvals									
Lead Market Size	5.05	4.85	5.68	5.59	9.37	4.81	5.03	5.11	4.45	5.32
	(2.19)	(2.33)	(2.25)	(2.38)	(4.22)	(2.26)	(2.30)	(2.22)	(1.38)	(2.30)
Lagged Dependent Variable	0.01									
	(0.22)									
R Squared	0.84	0.84	0.85	0.85	0.88	0.84	0.85	0.84	0.85	0.84
	Panel C: NLLS for Poisson model, dependent variable is count of generic drug approvals									
Market Size	0.97	11.69	12.60		18.45	13.77	12.45	11.86	6.32	12.22
	(6.27)	(3.50)	(3.30)		(6.48)	(3.78)	(3.40)	(3.23)	(2.05)	(3.28)
Lagged Dependent Variable	1.88									
	(0.44)									
R Squared	0.89	0.70	0.71		0.86	0.68	0.70	0.70	0.69	0.70

Notes: Huber-White standard errors in parentheses. Dependent variables are computed from the FDA dataset of New Drug Approvals. Panels A and C use current market size, and Panel B uses one period lead market size. Market Size is constructed as in Table II for columns 1-9 and 11. All regressions include drug and time dummies, use the income-based measure of market size and are weighted. In Panels A and C the number of observations is 198 in columns 2-5 and 8-10; and 186 in columns 6 and 7. The number of observations is 198 in Panel A column 1, 99 in Panel C column 1, 165 in columns 1-5 Panel B, and 8-10; and 155 in columns 6 and 7 Panel B. In column 1 the lagged dependent variable is instrumented with the twice lagged first difference of the dependent variable (see text). Life years lost is years prior to age 65 for each death in the U.S., calculated from the Mortality Detail Files (see text). Column 2 includes a count of total life years lost due to diseases in the corresponding category and time interval. Column 3 includes the amount of funding from NIH research grants in each category calculated from the CRISP database (see Appendix). 1940 /1970 trend is the log difference of drug approvals for category c between 1970 and 1940. In column 4, the 1940/1970 trend is interacted with period dummies (see text). Major drug category trends are linear time trends interacted with dummies for the 16 major drug categories (see text). In column 8, FDA approvals of biologics are added to the dependent variable. In column 10, market size includes information on health-care coverage (see text). In column 11, current market size is instrumented with the market size of the same cohort 5 years earlier.

APPENDIX TABLE A1

Summary of Disease Classifications and Drug Expenditure by Age Group

Class	Description	Expenditure Share x 1000 {Share of Expenditure by Age Group in Total Expenditure in Brackets} 0-20	20-30	30-50	50-60	60+	Age Group with Peak Expenditure
1	Antibiotics	1.20 {0.31}	0.58 {0.10}	0.62 {0.29}	0.62 {0.12}	0.90 {0.19}	0-20
2	Antivirals	0.04 {0.04}	0.02 {0.01}	0.37 {0.65}	0.34 {0.25}	0.05 {0.04}	30-50
3	Antiparasitics	0.01 {0.03}	0.03 {0.07}	0.06 {0.46}	0.04 {0.14}	0.08 {0.29}	30-50
4	Antifungals	0.28 {0.21}	0.22 {0.11}	0.22 {0.30}	0.24 {0.14}	0.38 {0.24}	30-50
5	Anemia	0.00 {0.04}	0.00 {0.02}	0.00 {0.19}	0.01 {0.27}	0.01 {0.47}	60+
6	Anticoagulants	0.01 {0.01}	0.00 {0.00}	0.03 {0.08}	0.12 {0.11}	0.77 {0.80}	60+
7	Glaucoma	0.00 {0.00}	0.01 {0.01}	0.02 {0.06}	0.07 {0.09}	0.58 {0.85}	60+
8	Acid/Peptic Disorders	0.12 {0.02}	0.25 {0.03}	0.72 {0.26}	1.28 {0.19}	2.87 {0.49}	60+
9	Antidiarrheals, Laxatives	0.00 {0.03}	0.00 {0.04}	0.01 {0.28}	0.01 {0.17}	0.03 {0.49}	60+
10	Cardiac	0.02 {0.00}	0.05 {0.01}	0.46 {0.14}	1.34 {0.17}	4.68 {0.67}	60+
11	Vascular	0.09 {0.01}	0.18 {0.01}	0.75 {0.15}	2.39 {0.19}	7.00 {0.64}	60+
12	Sedatives/Hypnotics, Antianxiety	0.02 {0.02}	0.08 {0.04}	0.24 {0.37}	0.28 {0.17}	0.58 {0.40}	60+
13	Antipsychotics/Antimanics, Antideppresants	0.31 {0.05}	0.69 {0.07}	1.71 {0.51}	1.48 {0.18}	1.28 {0.18}	30-50
14	Anorexiants/CNS Stimulants	0.12 {0.46}	0.02 {0.06}	0.05 {0.38}	0.02 {0.07}	0.01 {0.03}	0-20
15	Vitamins/Minerals	0.00 {0.04}	0.00 {0.02}	0.01 {0.19}	0.01 {0.16}	0.05 {0.58}	60+
16	Electrolyte Replenishment/Regulation, Water Balance	0.01 {0.02}	0.01 {0.01}	0.03 {0.11}	0.11 {0.15}	0.46 {0.71}	60+
17	Adrenal Corticosteroids	0.07 {0.22}	0.02 {0.05}	0.05 {0.25}	0.07 {0.15}	0.14 {0.34}	60+

Appedix Table A1 (cont.)

Class	Description	Expenditure Share x 1000 {Share of Expenditure by Age Group in Total Expenditure in Brackets} 0-20	20-30	30-50	50-60	60+	Age Group with Peak Expenditure
18	Androgens/Anabolic Steroids	0.00 {0.01}	0.00 {0.02}	0.01 {0.16}	0.00 {0.04}	0.07 {0.77}	60+
19	Estrogens/Progestins	0.06 {0.02}	0.70 {0.15}	0.48 {0.28}	1.27 {0.30}	0.97 {0.26}	50-60
20	Contraceptives	0.02 {0.05}	0.26 {0.42}	0.11 {0.50}	0.01 {0.03}	0.00 {0.01}	30-50
21	Blood Glucose Regulators, Thyroid/Antithyroid	0.04 {0.01}	0.14 {0.02}	0.50 {0.20}	1.36 {0.23}	2.90 {0.54}	60+
22	Topical Steroids	0.02 {0.14}	0.01 {0.06}	0.02 {0.24}	0.02 {0.12}	0.06 {0.43}	60+
23	Topical Anti-Infectives	0.01 {0.24}	0.01 {0.08}	0.01 {0.25}	0.01 {0.16}	0.02 {0.27}	60+
24	Extrapyramidal Movement Disorders	0.00 {0.01}	0.00 {0.01}	0.02 {0.14}	0.05 {0.11}	0.29 {0.74}	60+
25	Skeletal Muscle Hyperactivity, Anticonvulsants	0.19 {0.10}	0.23 {0.08}	0.50 {0.49}	0.37 {0.15}	0.39 {0.18}	30-50
26	Oncolytics	0.08 {0.04}	0.40 {0.15}	0.27 {0.27}	0.35 {0.15}	0.84 {0.39}	60+
27	Ocular Anti-Infective/Anti-Inflammatory	0.06 {0.21}	0.04 {0.08}	0.03 {0.20}	0.05 {0.12}	0.14 {0.39}	60+
28	Topical Otics	0.02 {0.31}	0.01 {0.11}	0.01 {0.19}	0.01 {0.11}	0.02 {0.28}	0-20
29	Vertigo/Motion Sickness	0.02 {0.13}	0.01 {0.06}	0.02 {0.22}	0.03 {0.17}	0.08 {0.42}	60+
30	Pain Relief	0.17 {0.03}	0.43 {0.06}	0.99 {0.36}	1.28 {0.19}	2.09 {0.35}	30-50
31	Antiasthmatics/ Broncodilators	0.45 {0.17}	0.26 {0.06}	0.32 {0.21}	0.56 {0.16}	1.28 {0.40}	60+
32	Nasal Decongestants, Antitussives, Cold Remedies	0.04 {0.12}	0.06 {0.11}	0.08 {0.42}	0.07 {0.15}	0.08 {0.20}	30-50
33	Antihistamines, Inhalation/Nasal	0.36 {0.15}	0.37 {0.10}	0.48 {0.35}	0.60 {0.18}	0.64 {0.22}	30-50

Notes: All data from the MEPS, 1996-1998. Construction of the 33 categories is described in the Data Appendix. Each category includes the indicated FDA sub-categories. Expenditure share is the total expenditure on drugs in the category divided by the total income of people in that age group. Share of expenditure by age group is the fraction of total expenditure in the category accounted for by the age group. Shares of expenditure by category are calculated for 10-year age groups. Age group with peak expenditure is the broad age group with the greatest expenditure on the corresponding category.

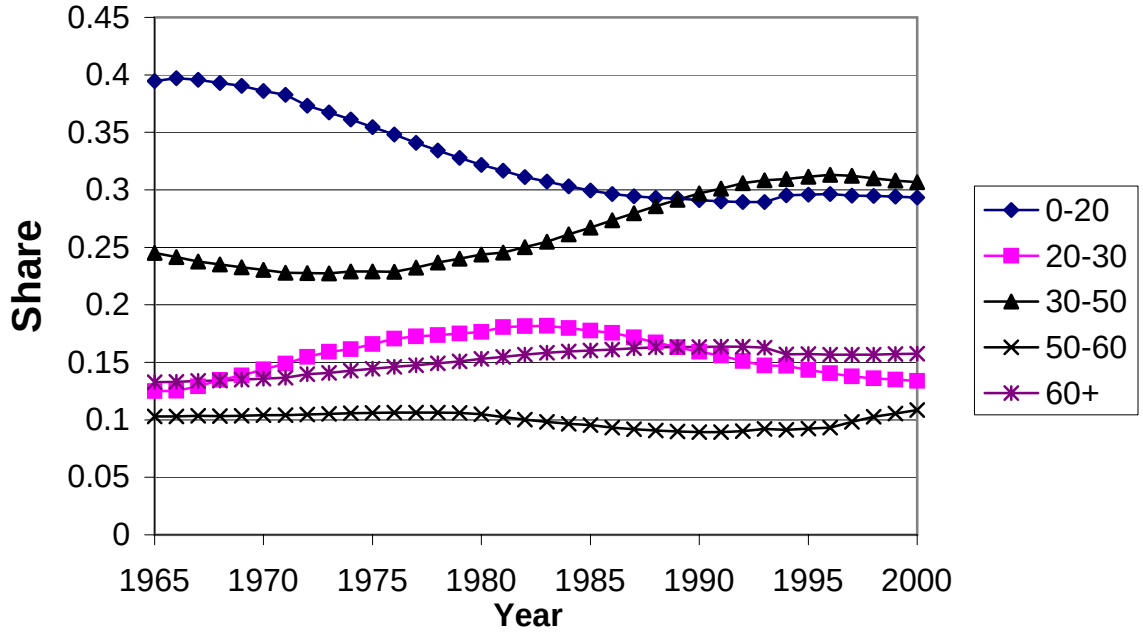
APPENDIX TABLE A2

Effect of Changes in Market Size, Using Previous Classification Scheme, and New Molecules and Patents as Dependent Variables

	(1)	(2)	(3)	(4)	(5)
	Panel A: NLLS for Poisson model, dep var is count of drug approvals				
Market Size	3.95	7.07		0.14	
	(1.16)	(3.78)		(2.96)	
Lag Market Size		-3.05	2.85		
		(3.17)	(1.00)		
Lead Market Size				6.26	6.41
				(3.83)	(1.61)
R Squared	0.87	0.87	0.87	0.91	0.91
	Panel B: NLLS for Poisson model, dep var is count of non-generic drug approvals				
Market Size	1.13	8.82		-0.49	
	(1.83)	(4.40)		(3.49)	
Lag Market Size		-6.89	-0.79		
		(3.48)	(1.63)		
Lead Market Size				6.83	6.31
				(4.40)	(2.18)
R Squared	0.81	0.81	0.81	0.85	0.85
	Panel C: NLLS for Poisson model, dep var is count of generic drug approvals				
Log Market Size	6.50	-0.72		13.74	
	(2.17)	(6.22)		(5.05)	
Lag Market Size		7.43	6.80		
		(5.51)	(1.67)		
Lead Market Size				-8.64	5.49
				(6.57)	(2.75)
R Squared	0.77	0.77	0.77	0.80	0.79
Number of Observations	204	204	204	170	170
	Panel D: NLLS for Poisson model, dep var is count of new molecules and patents				
Log Market Size	2.73		1.22		
	(1.97)		(1.72)		
Lead Market Size		4.35		0.38	4.60
		(2.18)		(2.16)	(1.85)
R Squared	0.72	0.63	0.87	0.94	0.94
Number of Observations	204	170	204	170	170
Dependent Variable	New Molecules	New Molecules	Patents	Patents	Patents
OECD Included	No	No	No	No	Yes

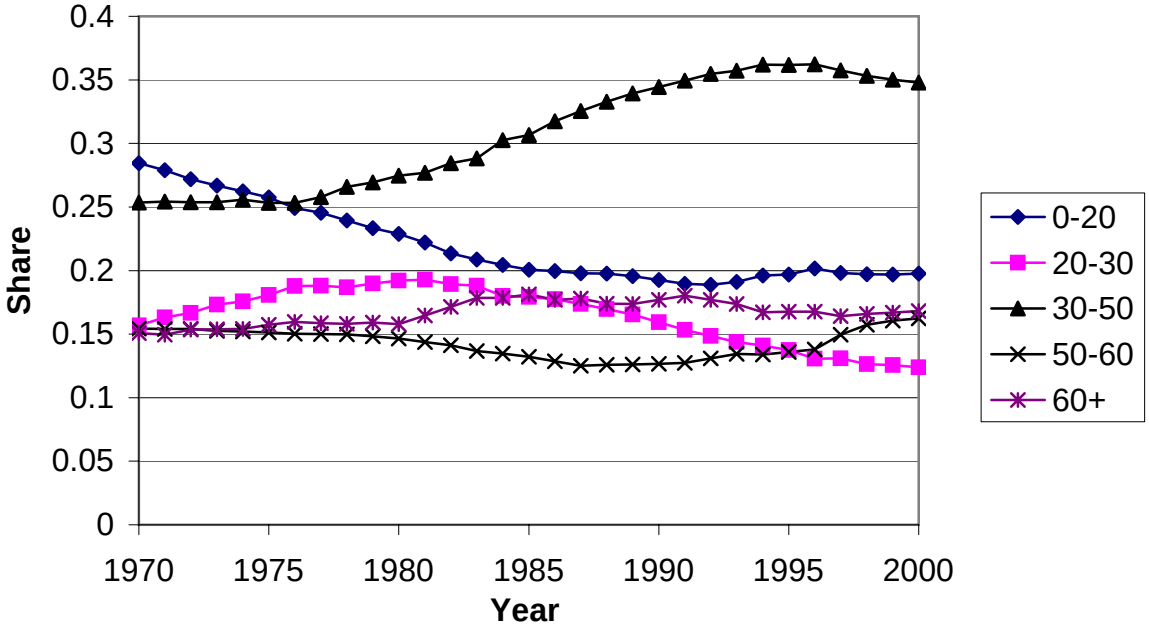
Notes: Huber-White robust standard errors are reported in parentheses. Variable construction and specifications are as described in Table II, except that the classification system from Acemoglu and Linn [2003] is used in Panels A, B and C. See text and Data Appendix for details. In Panel D, market size is income-based in columns 1-4, and population based, using OECD population, in column 5. The dependent variable in columns 1 and 2 in Panel D is a count of New Molecules in the FDA New Drug Approvals dataset. The dependent variable in columns 3-5 in Panel D is a count of patent, computed from the Derwent Inc. patent dataset (see Appendix).

Figure I
Share of Population by Age Group from CPS, 1964-2000



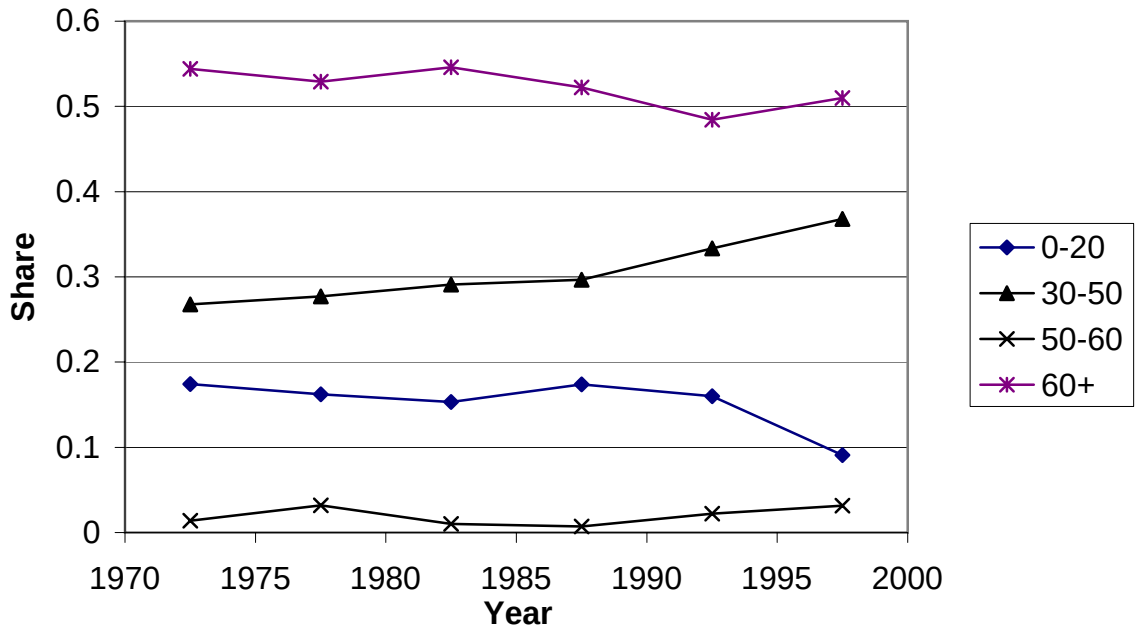
Notes: Share of population is the population of the corresponding age group divided by total population, computed from the March CPS.

Figure II
Share of Income by Age Group from CPS, 1970-2000



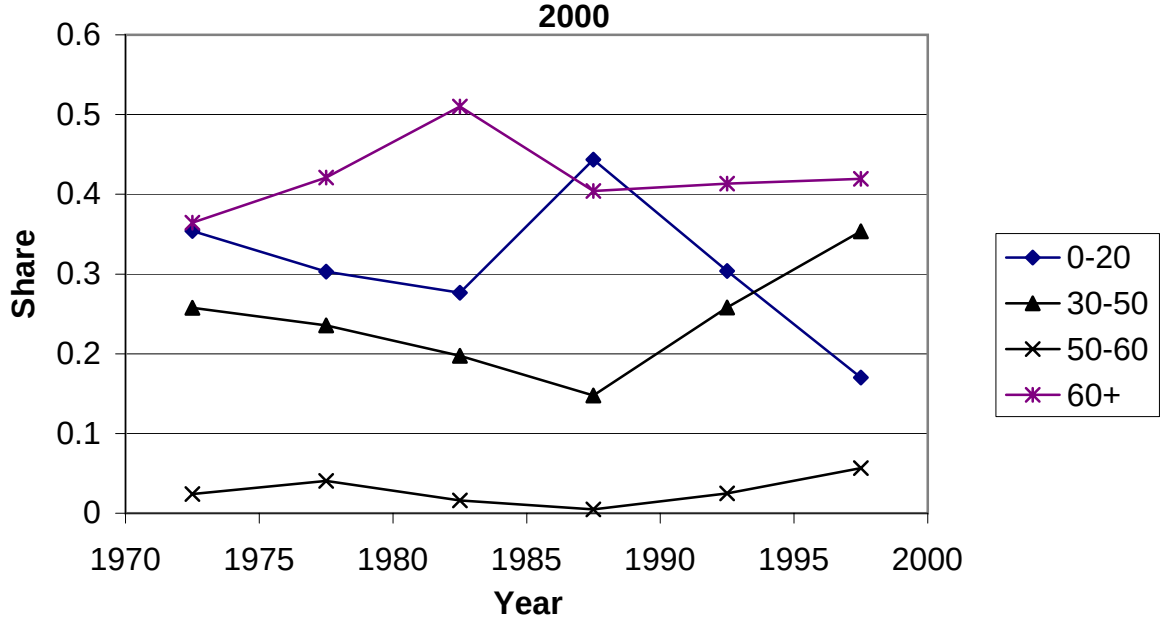
Notes: Share of income is income of the corresponding age group, divided by total income, computed from the March CPS. Individual income is obtained by dividing total family income equally among family members.

Figure III
Share of FDA Approvals by Age Group, 1970-2000



Notes: Share of FDA approvals is given by approvals of drugs in the corresponding broad age group divided by total approvals in that period, calculated from the FDA data set of New Drug Approvals. Each of our 33 drug categories is assigned to one of the five broad age groups according to which broad age group has the largest expenditure (see Appendix Table A1).

Figure IV
Share of Non-Generic FDA Approvals by Age Group, 1970-2000



Notes: Share of non-generic FDA approvals is given by approvals of non-generic drugs in the corresponding broad age group divided by total non-generic approvals in that period, calculated from the FDA data set of New Drug Approvals. Each of our 33 drug categories is assigned to one of the five broad age groups as in Figure III.