NBER WORKING PAPER SERIES

# THE DIFFUSION OF INNOVATIONS: A METHODOLOGICAL REAPPRAISAL

Manuel Trajtenberg

Shlomo Yitzhaki

Working Paper No. 1008

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge MA 02138

October 1982

NBER Working Paper #1008
October 1982

# THE DIFFUSION OF INNOVATIONS: A METHODOLOGICAL REAPPRAISAL

## Abstract

Studies of diffusion have traditionally relied on specific distributions -- primarily the logistic -- to characterize and estimate those processes. We argue here that such approach gives rise to serious problems of comparability and interpretation, and may result in large biases in the estimates of the parameters of interest. We propose instead the Gini's expected mean difference as a measure of diffusion speed, discuss its advantages over the traditional approach, and tackle with it the problems of truncated processes, inter-group comparisons, and related issues. We also elaborate on the use of the hazard rate, and suggest some possible extensions. The diffusion of CT scanners is presented as an illustration.

Manuel Trajtenberg
National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, MA 02138
Tel: (617) 868-3980

Shlomo Yitzhaki
Falk Institute of Economic Research
Hebrew University, Mt. Scopus
Jerusalem 91905

## 1. Introduction

One of the main difficulties hindering the study of technological change is the lack of good empirical counterparts to -- and hence of measures of -- many of the concepts that figure prominently in the theoretical analysis (e.g. "knowledge", "quality", "appropriability", etc.). The area of diffusion of innovations has been relatively fortunate in that respect, at least since the publication of Griliches' pathbreaking work on hybrid corn [10]. Indeed, much of the appeal of that paper stemmed from its having brought diffusion into the realm of the empirically measurable, i.e. from its having shown a way to quantify the phenomenon, and capture its essentials with the aid of just a few parameters (that could in turn be related to optimizing behavior). Yet, and notwithstanding the prolonged success of the paradigmatic approach established by [10], some basic methodological issues (primarily those associated with the reliance on specific distributions, most commonly the logistic) have not been satisfactorily resolved, and hence a reexamination is called forth. That is, for the most part, the task to be undertaken here, i.e. the dominant concerns in this paper are, once again, the definition of the concepts to be measured in diffusion, how to measure them, and how to illuminate and expand the analysis of the phenomenon with the aid of such measures. The focus is on "aggregate" diffusion, i.e., the pattern followed by the cumulative percentage of adapters over time, rather than on the underlying behavioral -- or "micro" -- aspects of the process (in Section 8, though, the relationship between the two levels of analysis is briefly discussed). Within that context we concentrate on diffusion speed, which is the parameter that has commanded the most attention in the literature.

Following a critical review of current methodology in Section 2, we pursue in Section 3 the prime objective of the paper, namely to put forward the Gini's expected mean difference ("Gini" hereafter) as a convenient measure of diffusion speed, elaborate on its meaning and implications, and outline its conceptual and methodological advantages over traditional measures. Section 4 deals with an intricate problem often encountered in diffusion studies: the measurement of diffusion speed of truncated processes. We develop a procedure based on the Gini that allows us to accurately measure the diffusion speed of just the observed segment of the process, and to compare processes that have been truncated at different levels, and/or estimated with different methods. Noting that it is often of interest to partition the universe of adopters into sub-groups, we present in Section 5 a way of assessing the impact of diffusion within each group on the diffusion speed of the aggregate process. In Section 6 we elaborate on (and advocate for) the use of the "hazard rate" as a simple yet incisive tool to probe further into the nature of diffusion processes.

The actual implementation of these ideas is illustrated in Section 7 via selected results from a case study on the diffusion of CT scanners -- a major innovation in medical technology -- in US hospitals. Finally, Section 8 sketches two possible extensions, linking diffusion with other topics in the economic and statistical literature.

To repeat, our interest throughout the paper centers on measurement issues, reflecting not only an obvious concern with "measuring things right", but also the belief that good measurement can breed good theory (the converse is usually taken for granted). In that vein, it seems that the replacement of prevailing ad hoc -- or distribution-specific -- estimates by a well-defined measure having general applicability, can only foster the much needed upgrading of diffusion theory.

## 2. A Critique of Current Methodology

Typically, empirical studies of diffusion of innovations have assumed that these processes follow a logistic pattern, i.e., that the cumulative distribution of adopters as a function of time is:[2,3]

$$(1) \qquad F(t) = 1 \,/\, [1 + \exp - (\alpha + \beta t)]$$

and proceeded to estimate the logit transform, using weighted or unweighted LS:

$$(2) \qquad \ln [F(t)/(1 - F(t)] = \hat{\alpha} + \hat{\beta} t$$

Attention is focused on the estimate $\hat{\beta}$ which, being the coefficient of time, is obviously related to some notion of speed of diffusion ( $\partial F(t)/\partial \beta > 0$ and hence the larger is $\beta$ the faster will be the diffusion process).[4] Different diffusion processes can thus be compared in terms of this coefficient, and often an attempt is made to explain the observed differences in diffusion speed by regressing the estimated $\hat{\beta}$'s on measures of profitability, size of investment, etc. In almost all such studies, LS estimates of equation (2)

---

2. Historically, the use of the logistic in diffusion studies is a clear case of methodological "spill-overs" from other disciplines, primarily from Biology (e.g., bio-assay) and Population Studies.

3. More generally, $F(t) = K/[1 + \exp - (\alpha + \beta t)]$ , where $0 < K \leq 1$ is the effective ceiling. We assume for the time being that $K = 1$ ; in Section 4 we discuss the case of $K < 1$ .

4. The intercept $\alpha$ has been, for the most part, dismissed as irrelevant, except for its role in setting an "origin", as defined in Griliches [10]. Whereas it is true that, being a constant of integration (to be precise, the constant of integration is $\exp \alpha$ ), $\alpha$ does not affect the shape of the distribution, it does convey an important piece of information: it can easily be shown that $-\alpha/\beta$ is the mean adoption time.

result in very high $R^2$'s (usually better than .90) and, notwithstanding the voicing of some reservations (e.g., Griliches [10, p. 505], Mansfield [16, p. 141]), that is taken as evidence upholding the logistic specification.

Now, if it could be safely assumed that most diffusion processes of interest do correspond to the logistic distribution (or, that even if they do not, that the resulting mispecification biases in the estimates were negligible), then no major objections could be raised against the received methodology. But, as we shall argue below, the fact is that there are neither compelling a priori reasons to sustain this particular specification, nor does the available empirical evidence provide strong support to it. The reliance on (2) and on the estimate $\hat{\beta}$ is therefore called into question.

As to theoretical considerations regarding the shape of aggregate diffusion processes:5 Diffusion has been until very recently one of those rare cases in Economics where, for better and for worse, the bulk of research was devoted to empirical studies rather than to theory. Thus, one can hardly expect to find theoretical results that will provide decisive support for any particular specification.

The conventional wisdom underlying most empirical studies had it that, due primarily to information-spreading and uncertainty-reducing mechanisms, the probability of adoption at any time $t$ would be related to the proportion of individuals that have already adopted by $t$ (this is often referred to as the "epidemic" or "contagion" effect). Or, more precisely, that the "hazard rate" (to borrow the term from Reliability Theory), defined as

---

5. A full discussion of the subject is beyond the scope of this paper -- here we shall content ourselves with sketching the main arguments.

$$(3) \qquad h(t) \equiv \frac{f(t)}{1 - F(t)} \quad , \quad f(t) = \frac{\partial F(t)}{\partial t}$$

would be a positive function of $F(t)$. In particular, it has been assumed that the relationship is linear, leading to the differential equation

$$(4) \qquad f(t) = \beta F(t)[1 - F(t)]$$

which solution is the logistic function (1). Notwithstanding some attempts to derive it more rigorously,[6] it is apparent that this amounts to no more than moving the assumption one step back. To be sure, (4) has sensible - if too vague - behavioral underpinings (i.e., the "contagion effect") whereas (2) by itself has none, but the fact remains that it is an ad hoc assumption rather than a result stemming from theoretical analysis.

Recently, though, various attempts have been made to model formally the behavioral phenomena underlying diffusion (see primarily Pakes [18], Jensen [12,13], and Reinganum [19]). The main ingredients of these models[7] and, for that matter, of any plausible theory of diffusion, are: (a) a model of decision-making regarding the adoption of the innovation by the individual; (b) the identification of those attributes of individuals that, in the context of

---

6. See, for example, Mansfield [16]: he postulates the general relationship $h(t) = \phi[F(t), \bar{x}]$, where $\bar{x}$ is a vector of explanatory variables (including primarily profitability and size of investment), approximates $\phi(.)$ with a Taylor expansion and, by dropping terms as necessary, arrives at an equation similar to (4), except that it includes the effect of $\bar{x}$ on $\beta$.

7. Except for Reinganum [19], where diffusion is the result of strategic behavior rather than differences among individuals. The problem there is that there is no mechanism to determine a unique ordering of adoption dates (as reflected in the fact that these are $n!$ equally likely Nash equilibria, $n$ being the number of adopters), which is precisely the question that a behavioral model of diffusion is supposed to answer.

(a), result in their having different responses at different times; (c) conjectures regarding the distribution of those attributes; (d) the specification of the "trigger mechanism", i.e., of the endogeneous or exogenous changes over time that activate the adoption decision. Clearly, this framework contains too many degrees of freedom for the ensuing theoretical analysis to be able to generate universal a priori restrictions on the shape of the aggregate diffusion process. Crucial in this respect is (c): As Pakes [18] has suggested, almost any diffusion curve can be arrived at by specifying suitable distributions of attributes. What theory can be expected to do is to provide a mapping of basic assumptions regarding (a) - (d) to broad families of diffusion patterns (e.g., S-shaped vs. concave, symmetric vs. asymmetric, etc.), but this remains to be done.

Now to the empirical evidence. To begin with, although being the most popular, the logistic has not been the only specification used in studies of diffusion. Examples to the contrary include: Bain [2], lognormal; Coleman et al. [6], both logistic and exponential;[8] Dixon [7], Gompertz; etc. More importantly, in those cases where the logistic was used, it is unwarranted to infer from the high $R^2$'s obtained that the logistic is in fact the "correct" specification. The econometric arguments stating the limitations of the $R^2$'s as a statistic to assess the functional form are well known and will not be repeated here. And, as was forcefully argued by Feller [8] long ago, there is an even more pervasive problem regarding the inferential value of goodness of fit measures of the logistic (or any other) distribution. To quote from Feller's later restatement [9, pp. 52-53]:

> ...an unbelievably huge literature tried to establish a transcendental "law of logistic growth"; measured in appropriate units, practically all growth processes were supposed to be represented by a [logistic] function....The only trouble with the theory is that not

---

8. Coleman, et al. did not estimate these functions but used them to represent, algebraically and graphically, different diffusion processes.

> only the logistic distribution, but also the normal, the Cauchy, and other distributions can be fitted to the same material with the same or better goodness of fit. In this competition the logistic distribution plays no distinguished role whatever; most contradictory theoretical models can be supported by the same observational material.

The crucial point is that the observed distribution will always fulfill, by necessity, the restrictions that characterize theoretical distributions, namely $0 \leq F(t) \leq 1$, and $F(t') \geq F(t)$ for all $t' > t$. Moreover, in most actual diffusion patterns studied, the data show strict concavity in its upper range and strict monotonicity all along, two of the properties exhibited by most distributions, and certainly by all those commonly used in this context. Thus, any functional form of that nature will offer a good fit and result, inter alia, in a high $R^2$. Yet, and contrary to what one may be led to believe from the forgoing discussion, the specification choice is by no means inconsequential for the issue of concern here, i.e., for the quality of the estimate of diffusion speed, provided of course that diffusion speed has been clearly defined in advance, independently of any particular distribution.

In order to gain some notion of the magnitude of the problem, we ran simulations in which the estimated equations were systematically misspecified, obtaining in all cases very good fits, but also large biases in the estimates of diffusion speed as it is defined in Section 3. The following is a typical example: Data were generated by an exponential distribution,

$$F(t) = 1 - e^{-\gamma t} \tag{5}$$

for various values of $\gamma$ and time lengths, but a logistic (the logit transform) was estimated instead; the $R^2$'s ranged from .95 to .99, but the biases in the estimates of diffusion speed were on the order of 50%.[9] Not surprisingly,

---

9. It is worth presenting explicitly one of these simulations: the distribution used to generate the observations was $F(t) = 1 - e^{-0.3t}$, $t \in [1, 25]$ in intervals of .5; the estimated equation was: logit = -0.62 + 0.33t, $r^2 = 0.99$. As shown below, $\hat{\beta}$ should be compared to $2\gamma$, i.e., 0.33 to .6, indicating a downward bias of almost 50%.

the Durbin-Watson statistics were very low (.12 - .18), reflecting the extent and nature of the misspecification.

Thus, the logistic can by no means be presumed to be the universal, or most accurate representation of diffusion processes and, if used indiscriminatingly, the estimates may be seriously biased. The alternative is to search in each case for the most appropriate specification, but then the question is how to compare between different diffusion processes, keeping in mind that an important goal of these studies is in fact to be able to make consistent and systematic comparisons. This poses a problem because nowhere in the literature is there to be found a well defined concept of speed of diffusion that could readily be computed for any distribution. As said, $\hat{\beta}$ - in (2) - is taken as the appropriate measure in the case of the logistic but, what are the equivalent parameters in, say, the normal or the log-normal distributions to which $\hat{\beta}$ should be compared? On the same note, studies using the logistic have not succeeded in providing $\hat{\beta}$ with a clear-cut meaning that will have conceptual and descriptive appeal (calling it the "rate" of diffusion is of no help): they only indicate a way of using $\hat{\beta}$ to calculate the time that the diffusion process takes to go from $F_1$ to $F_2$, these being two arbitrarily chosen points on the distribution.[10] It is rather unsatisfactory that this parameter, which had occupied such a prominent role in diffusion, cannot be grasped in its own terms and depends upon an element of arbitrariness for interpretation.

---

10. This is done as follows:

$$t(F_2) - t(F_1) = \frac{1}{\beta} \ln \left[\frac{F_2}{(1 - F_2)} \frac{(1 - F_1)}{F_1}\right] \equiv \frac{1}{\beta} \phi(F_1, F_2).$$

For example, if $F_2 = .8$ and $F_1 = .2$, $\phi(\cdot) = 2.77$; thus it can be stated that it took $2.77/\beta$ years for diffusion to go from 20% to 80%.

The issue of comparability arises vividly in Dixon [7]. He reexamined Griliches' study of hybrid corn and concluded, inter alia, that the appropriate specification for the majority of states was the Gompertz distribution, defined as

$$F(t) = a^{b^t}$$

rather than the logistic, as assumed by Griliches. He observes that ln b "performs a similar role to the b [ $\hat{\beta}$ in our notation] coefficient in the logistic function, in that it determines the rate at which $P_{it}$ [ F(t) in our notation] approaches the ceiling value" [7, p. 1456]. That is so but, more significantly in this context, the parameters $\hat{\beta}$ and ln b are in fact not exactly equivalent (that is shown in Section 3),[11] and therefore what is gained in precision (of the specification) is lost in comparability.

## 3. The Gini's Mean Difference as a Measure of Diffusion Speed

As stated in the Introduction, the main purpose of this paper is to propose the Gini's expected mean difference as a highly convenient summary parameter of diffusion processes. We contend that this statistic exhibits definite conceptual and methodological advantages over traditional measures, allowing it to overcome the difficulties described above, and to further extend the study of diffusion. The Gini is defined as[12]

---

11. Apparently, Dixon was aware of this fact, for he ran two separate regressions of diffusion parameters on profitability variables [7, Table IV], one for the states estimated with the Gompertz, and a second for those estimated with the logistic, whereas the preferred procedure (both from a conceptual and econometric viewpoint) would have been to run a single regression for all states.

12. Given that (6) is a double integral but $x_1$ and $x_2$ correspond in this case to the same variable t , it has to be divided by 2 so as to avoid "double counting".

$$(6) \qquad \Gamma = 1/2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} | x_1 - x_2 | \, f(x_1) \, f(x_2) \, dx_1 \, dx_2$$

where $x_1$ and $x_2$ are two independent, identically distributed random variables (or equivalently, two random realizations of the same variable $x$). In the context of diffusion the variable of interest is obviously time, $t_1$ and $t_2$ being the dates of any two adoptions. Writing (6) in a slightly different way,

$$(7) \qquad \Gamma = \int_{-\infty}^{\infty} \int_{t_1}^{\infty} (t_2 - t_1) \, f(t_2) \, f(t_1) \, dt_2 \, dt_1$$

its meaning becomes transparent: the Gini measures the expected time difference between any two adoptions over the whole diffusion process. Or to put it differently, it is the expected 'waiting time' of a random potential adopter at a random point in time (within the relevant time period).[13] This constitutes a well-defined notion of speed of diffusion,[14] and has generally applicability, i.e., is not contingent on specific distributions or any other pecularities of the processes studied.

Before turning to a more general measurement procedure, it is worth showing what the Gini is for those distributions commonly used in diffusion. This will not only help estimate the Gini in those cases where the underlying

---

13. In order to understand the precise meaning of the measure, it may be helpful to think of it as follows: a store announces a new product, to be sold on a first-come, first-serve basis, and all potential buyers (adopters) line up in a queue and wait for their turn. At regular time intervals a statistician calculates the expected waiting time for those still in the queue, and at the end computes the overall average, which is exactly the Gini.

14. "Speed" is commonly defined in terms of distance per unit time, except when the distance is a given, in which case the reciprocal (i.e., time per unit distance) is most often used (e.g., in foot races, or in describing the speed of a photographic camera). In diffusion the distance is indeed given by the unit interval, and hence it makes more sense to define speed in terms of $\Gamma$ rather than $1/\Gamma$. Moreover, the denominator of $\Gamma$ is not absolute distance but differences between any two points, making it hard to define speed as $1/\Gamma$.

distribution is known, but it also will allow to bring the results obtained in previous studies to a common denominator.

(i) The logistic distribution. A straightforward way to proceed in this case is to start up with (4) rather than (2), and to use a suitable transformation of (6),[15]

$$\Gamma = \int_{-\infty}^{\infty} F(t)\,[1 - F(t)]\,dt \tag{8}$$

Integrating (4),

$$\int_{-\infty}^{\infty} f(t)dt = \beta \int_{-\infty}^{\infty} F(t)\,[1 - F(t)]dt = \beta\Gamma \tag{9}$$

But the left hand of (9) equals 1, and therefore

$$\Gamma = \frac{1}{\beta} \tag{10}$$

(ii) The exponential distribution. Substituting (5) in (8),

$$\Gamma = \int_{0}^{\infty} e^{\gamma t}\,(1 - e^{-\gamma t})\,dt \quad \Rightarrow \tag{11}$$

$$\Gamma = 1/2\gamma \tag{12}$$

The Gini for the normal and log-normal are well-known, and that for the Gompertz is derived in the Appendix. Table I shows them in a condensed form.

In view of these results, we can now reexamine the problem in Dixon's work referred to above: the equivalent parameters there are not $\beta$ (from the logistic) and $\ln b$ (from the Gompertz), but $\beta$ and $\ln b/.7$ (more precisely, their reciprocals). The diffusion processes of <u>all</u> states can be thus compared, regardless of which distribution was fitted. Likewise, there is no longer need

---

15. This is done by substituting

$$\frac{1}{2}\,|\,t_1 - t_2\,| = \frac{t_1 + t_2}{2} - \min\,(t_1, t_2) \text{ in (6).}$$

See Kendall and Stuart [14] and Lomnicki [15].

Table I

The Gini for Various Distributions

| Distribution[a] | | | Gini |
|---|---|---|---|
| 1. | Logistic: | $F(t) = [1 + e^{-(\alpha+\beta t)}]^{-1}$ | $1/\beta$ |
| 2. | Exponential: | $F(t) = 1 - e^{-\gamma t}$ | $1/2\gamma$ |
| 3. | Normal:[b] | $t \sim N(\mu, \sigma^2)$ | $\sigma/\sqrt{\pi}$ |
| 4. | Lognormal:[c] | $\ln t \sim N(\mu, \sigma^2)$ | $e^{\mu+\sigma^2/2}\left[2\,\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1\right]$ |
| 5. | Gompertz: | $F(t) = a^{b^t}$ | $0.7/\ln b$ |

[a] For convenience, we avoid writing explicitly the cumulative distributions for the normal and lognormal.

[b] See Nair [17].

[c] See Aitchison and Brown [1]; $\Phi$ stands for the standard normal.

to split the states into two groups in order to assess the effect of profitability variables on interstate differences in diffusion speed, but a single regression having the Gini as dependent variable will do.

As we argued before, though, in most cases there is no good prior regarding the specific form of the underlying distribution, and the search for the correct specification can be cumbersome, and often inconclusive. In this respect as well the Gini proves to be highly advantageous, for it is possible to construct a simple, distribution-free measure of it. Integrating equation (8) by parts,

$$\Gamma = \int_{-\infty}^{\infty} F(t)\,[1 - F(t)]\,dt$$

define

$$u = t\ ,\ v = F(t)\,[1 - F(t)]\ ,\ v' = f(t)\,[1 - 2F(t)] \Rightarrow$$

$$\Gamma = tF(t)\,[1 - F(t)]\,\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} tf(t)\,[1 - 2F(t)]\,dt \tag{13}$$

But the first-hand term vanishes, and rearringing the second term,

$$\Gamma = 2\int_{-\infty}^{\infty} t[F(t) - 0.5]\,f(t)dt \tag{14}$$

Noting that $dF = f(t)dt$ , and changing variables accordingly,

$$\Gamma = 2\int_{0}^{1} t(F)(F - 0.5)\,dF = 2\ \mathrm{Cov}\,(t, F) \tag{15}$$

since $F$ distributes uniformily along the interval $[0,1]$ . If detailed information on each adoption is available, i.e., if the data consists of the vector $(t_1\ ,\ t_2\ ,\ \ldots,\ t_n)$ where $t_i$ is the adoption date of individual $i$ , then (15) can be computed simply by

(16) $$\hat{\Gamma} = 2 \text{ Cov } (\overline{F}, t)$$

where $\overline{F} = \frac{R_i}{n}$ , and $R_i$ is the rank of $t_i$ . But, the data are often aggregated in discrete time periods, in which case the integral in (15) has to be approximated linearly,

(17) $$\Gamma \cong 2 \sum_{i=1}^{n} \tilde{t}_i \, (\tilde{F}_i - 0.5) \, \Delta F_i$$

$$\tilde{t}_i = (t_i + t_{i+1})/2 \ , \ \tilde{F}_i = (F_i + F_{i+1})/2 \ , \ \Delta F_i = F_{i+1} - F_i$$

Thus, it is altogether unnecessary to resort to ad hoc assumptions or engage in specification search: the speed of diffusion, defined as the expected time difference between adoptions, can be computed readily from the original data by the covariance defined in (16), or by (17).

## 4. Estimating the Speed of Diffusion of Truncated Processes

We have assumed up to now that the diffusion process is observed in its entirety, i.e., that the data comprises the whole distribution $0 \leq F(t) \leq 1$. In many actual cases, though, $F(T) = p < 1$ , where $T$ is the last date for which data are available. Now, if it could be safely assumed that the process is near completion by $T$ , i.e., that $\lim_{t \to \infty} F(t) = K = p + \zeta$ , where $\zeta$ is a small fraction of $K$ , then it is possible to estimate $K$ (usually referred to as the effective ceiling) along with the other diffusion parameters. That necessarily entails the making of assumptions regarding the overall course of the process, most likely on the basis of its observed behavior up to $T$ , e.g., assuming a logistic distribution and estimating it using maximum likelihood or other

nonlinear methods.[16] The quality of the estimates so obtained will obviously depend on the validity of the behavioral assumptions (and thus be subjected to the same reservations rised above) and on how small $\zeta$ is.[17]

It is worth pointing out that, in fact, what this case implies is that the population of potential adopters was not correctly identified at the outset: the $(1-K)\%$ that did not and presumably will not adopt must have some distinctive characteristics that set them apart in terms of their behavior vis a vis the innovation. Thus, what is ultimately important is to identify those characteristics and delimit accordingly the "right" population set: a finding that $K < 1$ does not resolve the issue, only indicates that we have insufficient information.[18]

A more serious difficulty arises when there is no indication that the process is near completion by $T$ (that in turn implying that $p << 1$ and that there is no reason to believe that, if $K < 1$, $(K - p)$ is small) and no prior regarding the shape of the whole distribution.[19] Obviously, the only safe -- if

---

16. See for example Jarvis [11]; a further issue dealt with there is that $k$ is not necessarily constant, but may vary over time as a function of exogenous variables, such as prices.

17. It is important to note that the estimate of $\beta$ in the logistic is sensitive to $K$, and therefore the relative size of $\zeta$ is crucial for the precision of the former as much as for the latter.

18. This links back to the opening assumption "that the process is near completion": if that was actually known with a high degree of certainty -- such knowledge presumably stemming from having information on the relevant characteristics of the population -- it would imply knowing the approximate value of $K$ as well, and therefore its actual estimation can only improve its accuracy, but not render new information. If, on the other hand, there is not factual basis to assess the stage at which the process stands at $T$, then the estimate of $K$ can only be regarded as highly tentative, reinforcing the need to look closely at the attributes of the population.

19. As was already argued, the lack of good priors is the prevalent situation in most diffusion studies, but the problem is obviously aggravated when we observe a truncated distribution: the shorter is the range of the observed distribution, the less will be our ability to discriminate between alternative specifications, and hence the more arbitrary the choice becomes (and obviously the less reliable the estimates will be).

trivial -- solution is to wait until more data become available. But, this just evokes a basic tension that arises only too often in empirical research having a claim for "relevance": the longer the wait and hence the more complete and accurate the data, the more removed the study will turn out to be from current concerns, be them policy-oriented, or simply part of the quest to understand better the "here and now". And there certainly are plenty of concerns of that nature regarding diffusion, ranging from general issues such as: the extent to which the current productivity slowdown is related to a failure in the incentives to adopt innovations rather than to the drying up of inventive activity; whether there are structural differences between different sectors of the economy in that respect; to more specific, policy related issues such as the impact of government regulations on the diffusion of medical technologies. Moreover, the dilemma is made particularly acute in view of the long time span of most diffusion processes (10-30 years).

Clearly, any attempt to estimate the parameters of diffusion in these circumstances will render less than satisfactory results. The main contribution of the Gini in this respect is that it allows to accurately measure the speed of diffusion of the observed segment of the process, independently of any assumptions (implicit or explicit) regarding its future trajectory. Clearly, that is not possible if, instead, a particular distribution is assumed and estimated on the basis of the truncated distribution.[20] Partitioning the overall time span into two periods, the observed $(-\infty, T)$, and the unobserved $(T, \infty)$, (8) can be rewritten as:

---

20. One could assume that $K = p$ and estimate, say, the logistic; but that would entail an internal contradiction: if the true distribution -- having an unknown $K^* > p$ -- is really logistic, then it cannot be true that the truncated distribution also corresponds to a logistic with $K = p$, and thus the estimates will be necessarily biased.

$$(18) \qquad \Gamma = \int_{-\infty}^{T} F(t)[1-F(t)]dt + \int_{T}^{\infty} F(t)[1-F(t)]dt \equiv \Gamma^{o} + \Gamma^{u}$$

Integrating $\Gamma^{o}$ by parts as in (13):

$$(19) \qquad \Gamma^{o} = t\,F(t)[1-F(t)]\Big|_{-\infty}^{T} + 2\int_{-\infty}^{T} t[F(t) - 0.5]\,f(t)\,dt =$$

$$= p(1-p)T + 2p^{2}\int_{-\infty}^{T} t\left[\frac{F(t)}{p} - \frac{0.5}{p}\right]\frac{f(t)}{p}\,dt =$$

$$= p(1-p)T + 2p^{2}\int_{0}^{1} t(F^{*})(F^{*}-0.5)dF^{*} - p(1-p)\int_{-\infty}^{T} t\,f^{*}(t)\,dt$$

where $*$ indicates that the distribution has been normalized (i.e., $F^{*} = \frac{F}{p}$ and $f^{*} = \frac{f}{p}$), and $\bar{t} = \int_{-\infty}^{T} tf^{*}(t)dt$ is the average time of the observed period. Thus,[21]

$$(20) \qquad \Gamma^{o} = p^{2}\Gamma^{*} + p(1-p)(T-\bar{t})$$

where $\Gamma^{*} = 2\,\mathrm{Cov}(F^{*},t)$ is the Gini of the observed segment, calculated as if it were a complete process in itself (which is the implication of having normalized the distribution), i.e., it measures the speed of diffusion among those that adopted up to $T$, ignoring the fact that they are only a subset of the population of potential adopters. What $\Gamma^{o}$ does is to correct for that fact, thus measuring the contribution of the $p \times n$ initial adopters to the Gini of

---

21. More generally,

$$\Gamma^{o}(K) = \frac{1}{K^{2}}\left[p^{2}\Gamma^{*} + p(K-p)(T-\bar{t})\right]$$

given that $\Gamma^{*}$ is independent of $K$, this allows to easily compute $\Gamma^{o}$ under different assumptions regarding the ceiling. Note also that (20) is one of the many forms that the decomposition of the Gini can take: see, for example, Yitzhaki [26].

the entire process. This is as much as the data can tell without imposing additional structure on it and, as shown below, there are a variety of ways in which these partial measures can be used for comparative purposes. However, if there exists additional information that allows to form priors regarding the unobserved segment of the process (i.e., regarding $\Gamma^u$), then the Gini of the entire process, $\Gamma$, can be readily obtained by simply adding the prior to $\Gamma^o$. The advantage of this procedure over the fitting of a particular distribution to the whole process, is that it keeps observed phenomena strictly separate from conjectures (or projections), thus allowing to ascertain in a straightforward manner the effect on $\Gamma$ of different sets of assumptions regarding the remaining diffusion path, without these distorting the measure of the observed segment.

An assumption often made in these circumstances is that the process will exhibit in the future the same behavior -- on average -- as it did up to $T$ (we call this the "uniformity assumption") or, more precisely, that

$$(21) \qquad \frac{1}{F(T')} \int_{-\infty}^{T'} F(t)[1-F(t)]dt = \Gamma \quad \text{for all} \quad T' \geq T$$

It is easy to show that this property holds for the logistic, but for our purposes we need not assume that $F(t)$ corresponds to that distribution over its entire range, only that (21) holds on average over the period <u>following</u> $T$. The estimate of $\Gamma$ under this assumption is simply:

$$(22) \qquad \hat{\Gamma} = \frac{\Gamma^o}{p} = p\Gamma^* + (1-p)(T-\bar{t})$$

Note that (21) and (22) have important implications for comparing the partial measure $\Gamma^o$ to the existing body of results from past diffusion studies, which

constitutes the only readily available and natural reference group. Given that in most cases the original data used in those studies are not available, but only the published estimates of the diffusion parameters (and therefore $\Gamma^o$ cannot be computed for them), and that in all of them the estimates refer to whole prosesses (either because the data were indeed complete or because the truncation problem was assumed away), it is imperative to bring these estimates and $\Gamma^o$ to common grounds. This involves using the published estimates to evaluate the integral $\int_{-\infty}^{T'} F(t)[1-F(t)]dt$ , $T'=T(p)$ . As suggested above, if the estimated distribution was a logistic, then this integral is simply $\frac{p}{\beta}$ (it is $\frac{p^2}{2\gamma}$ in the case of an exponential distribution), but other distributions require that the integral be evaluated numerically. Thus, $\Gamma^o$ can be compared to, say, $\frac{p}{\beta}$ without this requiring any assumptions regarding $\Gamma^u$ . This simple result greatly facilitates the required comparisons, more so in view of the fact that most previous studies did in fact use the logistic.

Noting that comparing $\Gamma^o$ to $\frac{p}{\beta}$ is formally equivalent to comparing $\frac{\Gamma^o}{p} = \hat{\Gamma}$ to $\frac{1}{\beta}$ , it could be argued that there is in fact no way of escaping the uniformity assumption. But this is not so: the comparison of $\Gamma^o$ to $\frac{1}{\beta}$ places the uniformity assumption (i.e., "the burden of the proof") on the <u>other</u> process (the one which $\frac{1}{\beta}$ corresponds to), and that represents no extra restriction: the assumption, justified or not, was there to begin with, implied in the fitting of a logistic distribution.

A different case arises when original data (i.e., the vectors $F(t),t$) for all the processes to be compared are actually available. This is the likely situation when the diffusion of a particular innovation is being studied, but the total population of adopters is divided into subsets, each generating its own process, and the objective is to compare between them (e.g., Griliches' study of hybrid corn, by geographical areas). Suppose that there are $m$ such

processes, and that they have reached levels $p_1$ , $p_2$ , ..., $p_m$ ( $p_i \gtreqless p_j$ ) by T . As before, if reasonable conjectures can be made on the $I^u_i$'s , then the processes can be compared in terms of their estimated $I_i$'s ( $I^o_i + I^u_i$ ) , where the $I^o_i$'s are calculated using all the data available for each process. Otherwise, the processes have to be truncated at the same cut-off level: $p_0 = \underset{i}{\text{Min}} (p_1, p_2, .., p_m)$, and the $I^o_i$'s (in which terms the comparisons are to be made)[22] are calculated using only the $p_0 \times n_i$ initial observations of each process. This entails loosing information at the upper end of those processes with relative high $p_i$'s , which is the cost to be paid for not resorting to assumptions regarding the unobserved segment of the process. There is no way to avoid this trade-off, and no dominant strategy: the course of action to be taken will depend upon the particulars of each study.

It is important to stress again the partial and hence tentative nature of all these comparisons: as more data become available the measures ought to be revised and the comparisons redone.

Finally, it should be clear that the procedures described here apply to any case in which some part of the distribution is missing, and not just its upper end. In Russell [20] for example, the data on the initial stages of two of the innovations studies are severely lacking (one starts at 19% and the other at 48%), which is not an uncommon situation: data on diffusion are often gathered only after the innovation has become important enough and hence widely spread.

---

[22]Note that truncating the process at a common level $p_0$ usually implies $T_i \neq T_j$ and $\bar{t}_i \neq \bar{t}_j$ , and that these are in fact parameters of interest in themselves. Now, if $(T_i - \bar{t}_i) \cong (T_j - \bar{t}_j)$, i,j = 1, ..., m , then $I^o_i = p_o^2 I^*_i + c$ , where $c = p_0(1 - p_0)(T - \bar{t})$ , and hence there is no need to compute $I^o_i$ , but $I^*_i$ -- which enjoys some advantage in interpretation -- suffices for comparative purposes.

## 5. Decomposition by Groups

As mentioned in Section 4, it is often of interest to divide the population of adopters into sub-groups, and do a comparative study of their diffusion processes. A related issue is to investigate the impact of each group on the aggregate process, i.e., the extent to which diffusion within each group accelerates or slows down overall diffusion. Formally, this involves decomposing the "overall" Gini into group-specific components that capture the relative size of each group, and the similarity of its process to (or its correlation with) the aggregate process.

Let $F_i(t)$ be the diffusion process and $n_i$ the size of group $i$, $i = 1, \ldots, m$. The aggregate process will then be

$$F_o(t) = \frac{1}{N} \sum_{i=1}^{m} n_i F_i(t) , \qquad N = \sum_{i=1}^{m} n_i \tag{23}$$

and the overall Gini,

$$\Gamma_o = \int_{-\infty}^{\infty} [1-F_o(t)]F_o(t)dt \tag{24}$$

Substituting (23) for $F_o$ in (24),

$$\Gamma_o = \sum_{i=1}^{m} \Gamma_{o.i} \tag{25}$$

where

$$\Gamma_{o.i} = \frac{ni}{N} \int_{-\infty}^{\infty} [1-F_o(t)]F_i(t)dt , \qquad i = 1, \ldots, m \tag{26}$$

The $\Gamma_{o.i}$'s are the magnitudes of interest: the larger the correlation between the diffusion process in group $i$ and the aggregate process (i.e., the larger is the integral in (26)), the larger $\Gamma_{o.i}$ is, and hence the more group $i$ will <u>slow down</u> aggregate diffusion (the same holds, <u>mutatis mutandis</u>,

for the relative size of group $i$, $n_i/N$). For comparative purposes, though, it is more convenient to use the shares:

$$(27) \qquad w_i = \Gamma_{o.i}/\Gamma_o \,, \qquad \sum_{i=1}^{m} w_i = 1$$

which meaning is immediate: $w_i$ is simply the fraction of the overall Gini accounted for by group $i$ or, in other words, it is the percentage contribution of the diffusion process in group $i$ to the average waiting time between adoptions in the total population.

The actual computation of (27) is done as follows: let $t_o$ be the vector of adoption times of the aggregate process, and $t_i$ the analogous vector for group $i$. Define a new vector $\tilde{t}_i$ for each group $i$, so that its $j^{th}$ element is:

$$\tilde{t}_{ij} = \begin{cases} t_{oj} & \text{if } t_{oj} \in t_i \\ 0 & \text{if } t_{oj} \notin t_i \end{cases}$$

thus, $t_o = \sum_{i=1}^{m} \tilde{t}_i$, $\Gamma_o = 2\,\mathrm{Cov}\,(F_o, t_o) = 2 \sum_{i=1}^{m} \mathrm{Cov}\,(F_o, \tilde{t}_i)$,

and

$$(28) \qquad w_i = \mathrm{Cov}\,(F_o, \tilde{t}_i)/\mathrm{Cov}\,(F_o, t_o)$$

Finally, it is worth noting that the decomposition here is formally similar to the one performed in the context of the familiar CAPM: the groups in diffusion can be thought of as different stocks, and aggregate diffusion as the market portfolio. As shown in Shalit and Yitzhaki [23], the $\Gamma_{o.i}$'s -- properly normalized -- are analogous to the $\beta$'s in CAPM, a fact that facilitates the interpretation of these measures, and may prove useful to further explore the links between diffusions by groups and aggregate diffusion.

## 6. The Adoption Rate

So far we have been concerned exclusively with the measurement of the speed of diffusion, as a one-parameter representation of the diffusion process. The next question is whether it is possible to learn more about the process itself, still within the same restrictive framework, i.e., having data only on adoption dates (or percentages of adopters in discrete time periods), and without imposing additional structure on it. The answer is a qualified yes: there are obviously numerous properties of the observed diffusion process that could be sought, and at least as many statistical tests that could be applied to them. But, apart from purely descriptive purposes, it is worth investigating a characteristic only in so far as it enhances the understanding of diffusion as a socio-economic phenomenon, or if it gives some indication for further research. Of course, it is theory's customary role to provide guidance in that respect but, as stated before, the results are wanting. Thus, any further step taken in this direction will be necessarily tentative, and no general conclusions can be expected.

We want to suggest the behavior over time of the hazard rate (which could be appropriately relabeled in this context the "adoption rate") as a likely candidate for investigation. To recall, the adoption rate is formally defined as[23]

$$h(t) = \frac{f(t)}{1 - F(t)}$$

By analogy to its meaning in Reliability Theory, $h(t)$ can be interpreted here as the conditional probability of adoption at $t$, having been a "hold-out" until

---

23. For a description of the properties of the hazard rate, see Barlow and Prochan [3]. The relationship between the hazard rate and the Gini is discussed in Chandra and Singpurwalla [5].

then. Now, if $h(t) = \gamma$, i.e., if the conditional probability is constant over time, then the underlying distribution is necessarily exponential, suggesting that the "contagion effect" (that is, the direct influence of previous adopters on would-be adopters) is not the predominant force driving the diffusion process. On the other hand, if $h'(t) > 0$, the corresponding distribution is likely to be S-shaped, and the contagion effect may be at work, but no solid inferences can be drawn without further information (the case of $h'(t) < 0$ does not seem to be relevant in diffusion).

A good example is provided in the study of Coleman et. al. [6] on the diffusion of the use of a new drug among physicians. They use the adoption rate (without referring to it as such) in order to distinguish between what they call "snowball" vs. "individual" diffusion processes. The most telling finding is that the adoption rate of socially integrated doctors increases sharply over time, whereas that of isolated doctors oscillates without displaying a trend. Thus, they conclude that the contagion effect -- associated with $h'(t) > 0$ -- is at work in the former case but not in the latter, for which $h'(t) = 0$.[24] Note that what allows the authors to draw these conclusions is not just the sign of $h'(t)$, but the fact that adopters were separated into subgroups according to variables reflecting the extent of their social integration, which has a direct bearing on the plausibility of the contagion effect.

The simplest procedure will thus be to regress $h(t)$ on $t$ (and/or on $F(t)$) and test for the significance of the slope coefficients. But, this test will be meaningful only if the diffusion process displays a uniform behavior over time. In some cases, though, the observed distribution results from the concatenation of different sequential processes, each initiated in response to

---

24. Following the same reasoning, they characterize the diffusion process among socially integrated doctors as logistic, and that of isolated doctors as exponential.

discrete changes in the exogenous variables governing the adoption decision (e.g., major technological improvements in the innovation, jumps in prices, changes in government policies, etc.). This sort of phenomena cannot be captured in the simple correlation between $h(t)$ and $t$, but will most likely show-up in a plot of $h(t)$ on $t$: as will become apparent in the empirical illustration below, the visual inspection of such plots can be highly informative, and provide the researcher with much needed guidance for measuring and analyzing the diffusion process.

## 7. An Empirical Illustration: The Diffusion of CT Scanners

The approach laid out in the preceeding sections can be best appraised by applying it to a concrete case, for which purpose we have chosen the diffusion of CT Scanners in U.S. hospitals. CT (Computed Tomography) is a revolutionary diagnostic technology that produces highly detailed and accurate pictures of thin "slices" of any section of the human body, using a sophisticated configuration of x-rays, detectors and computers. Developed in the late sixties, the first operational prototype was built by the British firm EMI in 1971, and the first commercial installation in the U.S. took place in June 1973. It has commanded a great deal of public attention ever since, partly because of its scientific merits, but also because of growing concerns that this kind of expensive advances in medical technologies may have been fueling the rapid rise in health care costs. Acting on this belief, the government enacted a series of regulations designed to slow-down the diffusion of CT scanners. It is still a matter of controversy whether diffusion was indeed "too fast", and whether those regulations have had a noticeable effect on it. Thus, the interest in the case is not only academic, but has policy implications as well.

The data used in this study consist of the adoption dates (month, year) of all first scanners (some medical institutions have acquired more than one) installed in community hospitals during the period 6/73 - 12/81,[25] and supplementary information from the American Hospital Association (AHA) Annual Surveys of Hospitals. Table II shows the distribution of hospitals and adopters by bed-size category. Given that only 1.2% of hospitals with less than 100 beds had CT scanners, we decided to exclude them from the study. Still, the diffusion process for community hospitals with beds > 100 (hereafter referred to just as "hospitals") is far from complete ($p = .39$) and therefore the methods developed in Section 4 for truncated distributions are called into action.

The previous statement implies having an idea of the value of $K$; and indeed, although it is hard to predict at this stage what will be the precise ceiling for CT scanners, it can be safely assumed that it will not surpass the ceiling reached by Diagnostic Radioisotopes, a previous innovation in imaging technology,[26] shown in Table III. Thus, most of the calculations will be done for $K = .84$, but lower ceilings will be considered as well ($K = .7$ and $K = .6$; it is highly implausible that $K$ will be less than 60%). We discuss below "overall" diffusion (i.e., diffusion among all hospitals), diffusion by bed-size groups, and by type of control.

---

25. These data were collected by the first author as part of a much wider research project on CT scanners. We report here only partial results that have a bearing on the methodological issues addressed in this paper.

26. This conjecture is based on information regarding the relative prices of the two systems, the availability of competing technologies along the diffusion path, etc.

Table II

Distribution of Community Hospitals[a] and Adopters of CT Scanners by Bed-Size

| | (1) | (2) | (3) |
|---|---|---|---|
| Number of Beds | Total Number of Hospitals | Hospitals with CT Scanners | p: (2) ÷ (1) |
| up to 99 | 2,848 | 34 | .012 |
| 100 - 199 | 1,417 | 187 | .132 |
| 200 - 299 | 719 | 285 | .396 |
| 300 - 399 | 384 | 254 | .662 |
| 400 - 499 | 244 | 201 | .824 |
| 500+ | 314 | 275 | .876 |
| TOTAL | 5,926 | 1,236 | .209 |
| 100+ | 3,078 | 1,202 | .390 |

[a]The definition of "community hospitals" used by the AHA is: "Non-federal, short-term general and other special hospitals, excluding hospital units of institutions". This excludes approximately 1,200 long-term and/or federal hospitals, of which only 37 had CT scanners.

Table III

Ceilings for Diagnostic Radioisotopes

| Bed-size: | 100-199 | 200-299 | 300+ | All |
|---|---|---|---|---|
| Ceiling: | 70% | 92% | 98% | 84% |

Source: AHA 1977 Survey.

7.1 Overall Diffusion

The objective here is to estimate the diffusion speed of CT scanners so as to compare it with other innovations. To recall, if we want to avoid making assumptions about the future course of the process, the comparisons are to be based on the partial measure $\Gamma^o$ (see equation (20)) which, allowing for different K's can be written as:

$$(29) \qquad \Gamma^o\left(\frac{p}{K}\right) = \frac{p}{K}\left[\frac{p}{K}\,\Gamma^o + \left(1 - \frac{p}{K}\right)(T - \bar{t})\right]$$

The availability of very detailed data allowed us to compute $\Gamma^*$ using (16) rather than (17) (i.e., using ranks), rendering a value of $\Gamma^* = 12.89$, i.e., the average time difference between any two adoptions for the set of hospitals that adopted CT scanners up to $T = 12/81$ was of slightly more than a year.[27] Plugging it in (29), $\Gamma^o$ was computed for the various K's, and the results are presented in Table IV.

Table IV

Estimates of $\Gamma^o$

| $\Gamma^o(K)$ \ K | .84 | .7 | .6 |
|---|---|---|---|
| in months | 14.06 | 15.20 | 15.77 |
| in years | 1.17 | 1.27 | 1.31 |

27. The time unit of the data and hence of the estimates is months ( t is the number of months elapsed since Nov. 1972, the date when the innovation was first announced and displayed in the U.S.). For some comparisons, though, the estimate will be transformed into years.

We want now to compare these with estimates of the diffusion speed of 20 innovations reported in the literature (see Table V). All of them were estimated using the logistic specification and, as shown in Section 4, the tranformation of the reported $\hat{\beta}$'s that allows comparability with our results is simply $f^{o}\left(\frac{p}{K}\right) = \frac{p/K}{\beta}$. Figure 1 summarizes the results: The diffusion of CT scanners was indeed quite fast, i.e., it belongs to the fastest third of the innovations studied. Moreover, it was more than 3 times faster than the diffusion of the other two reported innovations in diagnostic technologies (electroencephalograph and diagnostic radioisotopes). As Figure 1 makes clear, the conclusions hold for the three alternative ceilings considered.

## 7.2 The Adoption Rate

We examine now the bahavior of the adoption rate which, as argued in Section 6, may provide some insight into the nature of the process. Regressing it on time and on F(t),[28]

$$h = \underset{(3.5)}{0.003} + \underset{(4.9)}{0.000058}\, t \qquad r^2 = .20 \tag{30}$$

$$h = \underset{(7)}{0.004} + \underset{(4.7)}{0.01}\, F(t) \qquad r^2 = .19 \tag{31}$$

These results are informative only in a "negative" sense: (31) makes it highly unlikely that the process as a whole corresponds to the logistic, because (a) the very low $r^2$, which indicates that $h(t)$ was not smoothly increasing with F(t), as it should have been if the underlying distribution was a logistic, and (b) the very small coefficient of F(t) which, appropriately transformed, renders an estimate of diffusion speed 3 times larger (i.e., 3 times slower) than that calculated with the Gini in Table IV (for K = .84). Likewise, (30) rules out the possibility that the process conforms to an exponential-type

28. These calculations have been done for K = .84.

Table V

$\hat{\beta}$ and $\Gamma^{o}(K)$ for Twenty Innovations[a]

| Innovation | (1) Reported $\hat{\beta}$ | (2) $\Gamma^{o}(K=.84)$ | (3) $\Gamma^{o}(K=.7)$ | (4) $\Gamma^{o}(K=.6)$ |
|---|---|---|---|---|
| Griliches [10] | | | | |
| 1. Hybrid corn[b] | 0.54 | 0.86 | 1.03 | 1.20 |
| Jarvis [11] | | | | |
| 2. Improved pastures in Uruguay[c] | 0.55 | 0.84 | 1.01 | 1.18 |
| Mansfield [16] | | | | |
| Bituminous coal mining: | | | | |
| 3. Shuttle Car | 0.32 | 1.45 | 1.74 | 2.03 |
| 4. Trackless mobile loader | 0.32 | 1.45 | 1.74 | 2.03 |
| 5. Continuous mining loader | 0.49 | 0.95 | 1.14 | 1.33 |
| Iron and Steel: | | | | |
| 6. By-product coke oven | 0.17 | 2.73 | 3.28 | 3.82 |
| 7. Continuous wide-strip mill | 0.34 | 1.37 | 1.64 | 1.91 |
| 8. Continuous annealing line for tin plate | 0.17 | 2.73 | 3.28 | 3.82 |
| Brewing: | | | | |
| 9. Pallet loading machine | 0.55 | 0.84 | 1.01 | 1.18 |
| 10. Tin container | 2.40 | 0.19 | 0.23 | 0.27 |
| 11. High-speed bottle filler | 0.36 | 1.29 | 1.55 | 1.81 |
| Railroads: | | | | |
| 12. Diesel locomotive | 0.20 | 2.32 | 2.79 | 3.25 |
| 13. Centralized traffic control | 0.19 | 2.44 | 2.93 | 3.42 |
| 14. Car retarders | 0.11 | 4.22 | 5.06 | 5.91 |
| Romeo [22] | | | | |
| 15. Numerically controlled machine-tools[b] | 0.35 | 1.33 | 1.59 | 1.86 |
| Russell [20][b] | | | | |
| 16. Post operative recovery room | 0.31 | 1.50 | 1.80 | 2.10 |
| 17. Intensive care unit | 0.30 | 1.55 | 1.86 | 2.17 |
| 18. Electroencephalograph | 0.12 | 3.87 | 4.64 | 5.42 |
| 19. Diagnostic radioisotopes | 0.13 | 3.57 | 4.29 | 5.00 |
| 20. Respiratory therapy | 0.41 | 1.13 | 1.36 | 1.59 |

Table V, continued

Footnotes

[a]The $\hat{\beta}$'s are the logistic coefficients reported in the studies listed here (the time unit is years). In columns (2) to (4), the $\hat{\beta}$'s are transformed into the partial measure $\Gamma^{o}(p/K)$ for different K's, so as to enable comparisons with CT scanners. To recall, if F is logistic then:

$$\Gamma^{o}(p/K) = \int_{-\infty}^{T'} F(1-F)dt = p/K\hat{\beta} \quad , \qquad T' = T(p/K)$$

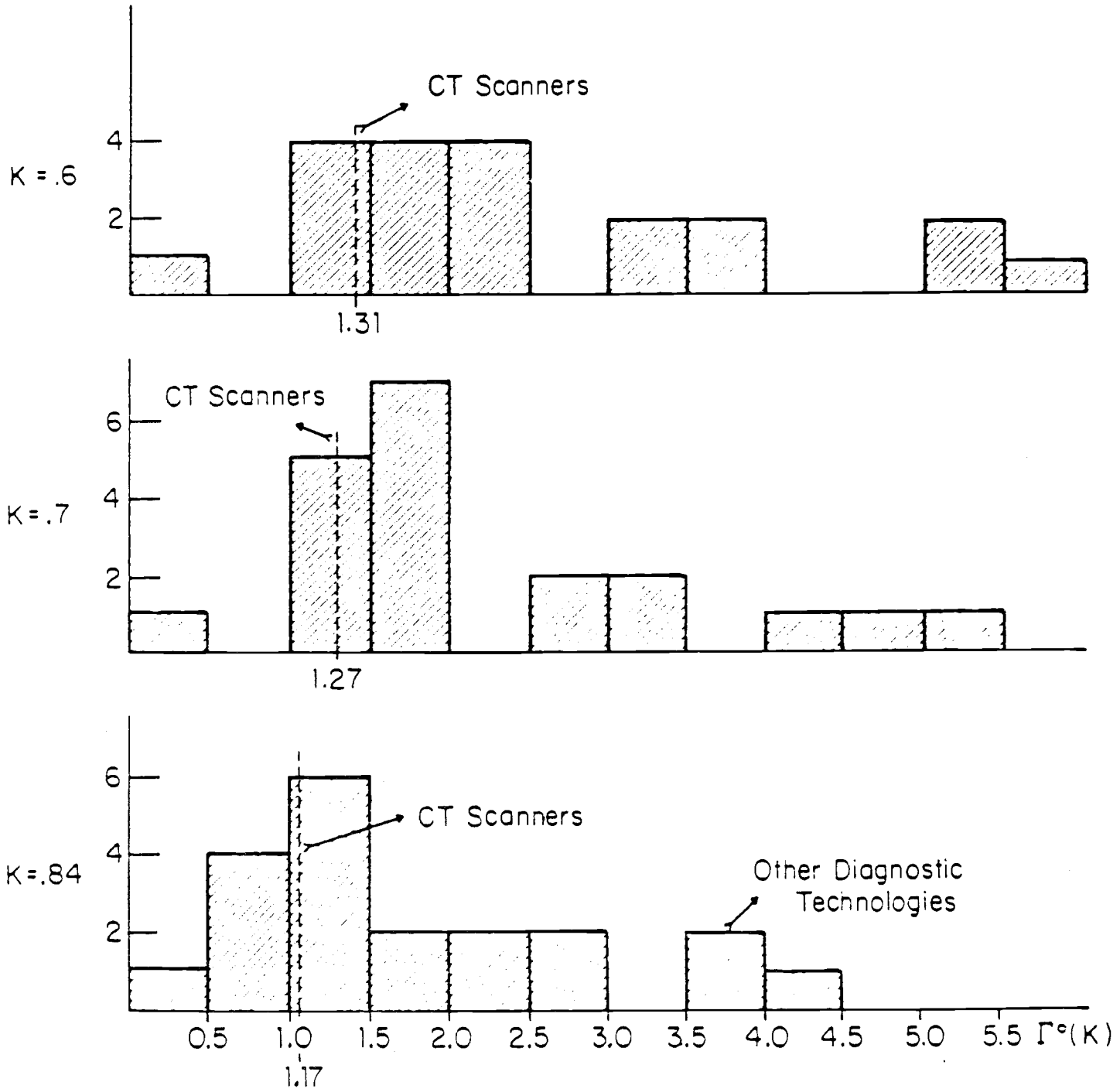
According to Table II, $p = .39$, hence $\Gamma^{o} = .39/K\hat{\beta}$.

[b]In all these cases we have taken the average of the group coefficients: states in Griliches, industries in Romeo, and different classes of hospitals in Russell (the estimates for hospitals with beds < 100 were excluded). This is not equivalent to the coefficient of the aggregate process, but it is a good enough approximation for our purposes.

[c]Jarvis tried several specifications, but the coefficients did not vary much: the range was .51 - .59 with an average of .55, which is the figure reported here.

Figure 1

The Frequency Distribution of $\Gamma^{\circ}(K)$ for 20 Innovations




Sources: Table 4, and Table 5, Columns (2)-(4)

(30) rules out the possibility that the process conforms to an exponential-type distribution, because the time coefficient, although very small, is nevertheless significant. Thus, a more complex pattern is suggested instead. The plot of $h(t)$ on time (see Figure 2) clarifies the issue: there was a sharp discontinuity midway along the diffusion path (in the third quarter of 1977), the process behaving very differently before and after. In the first period $h'(t) >> 0$, suggesting an S-shaped distribution (and the "contagion effect"), whereas afterwards there is a drop in the level of $h(t)$, and $h'(t) \cong 0$, implying a slow-down in the process and an exponential-like pattern thereafter. Following this lead we computed separate $I^o$'s for each period, and found that the first was faster than the second by a factor of 2. Furthermore, the logistic fits well the first period, rendering an estimate of diffusion speed consistent with the partial Gini. A discussion of the causes underlying this rather dramatic change is beyond the scope of this paper; suffice it to say here that it was due primarily to the implementation of government regualtions, and prior expectations in this regard.

The important point is that, had we proceeded according to the received methodology, we would have probably overlooked this crucial feature of the diffusion process (as can be seen in Figure 3, the plot of $F(t)$ -- often shown in diffusion studies -- does not reveal it either) and assuming a particular distribution would have resulted in biased estimates. To illustrate, we estimated the logistic for the whole period (and different K's). Comparing the results in Table VI to those in Table IV, the logistic overestimates the speed of diffusion by 50-80%.[29]

---

29. According to the logistic estimates, CT scanners would have been the second fastest innovation, the first being tin containers in Mansfield's study (which is a puzzling outlier).

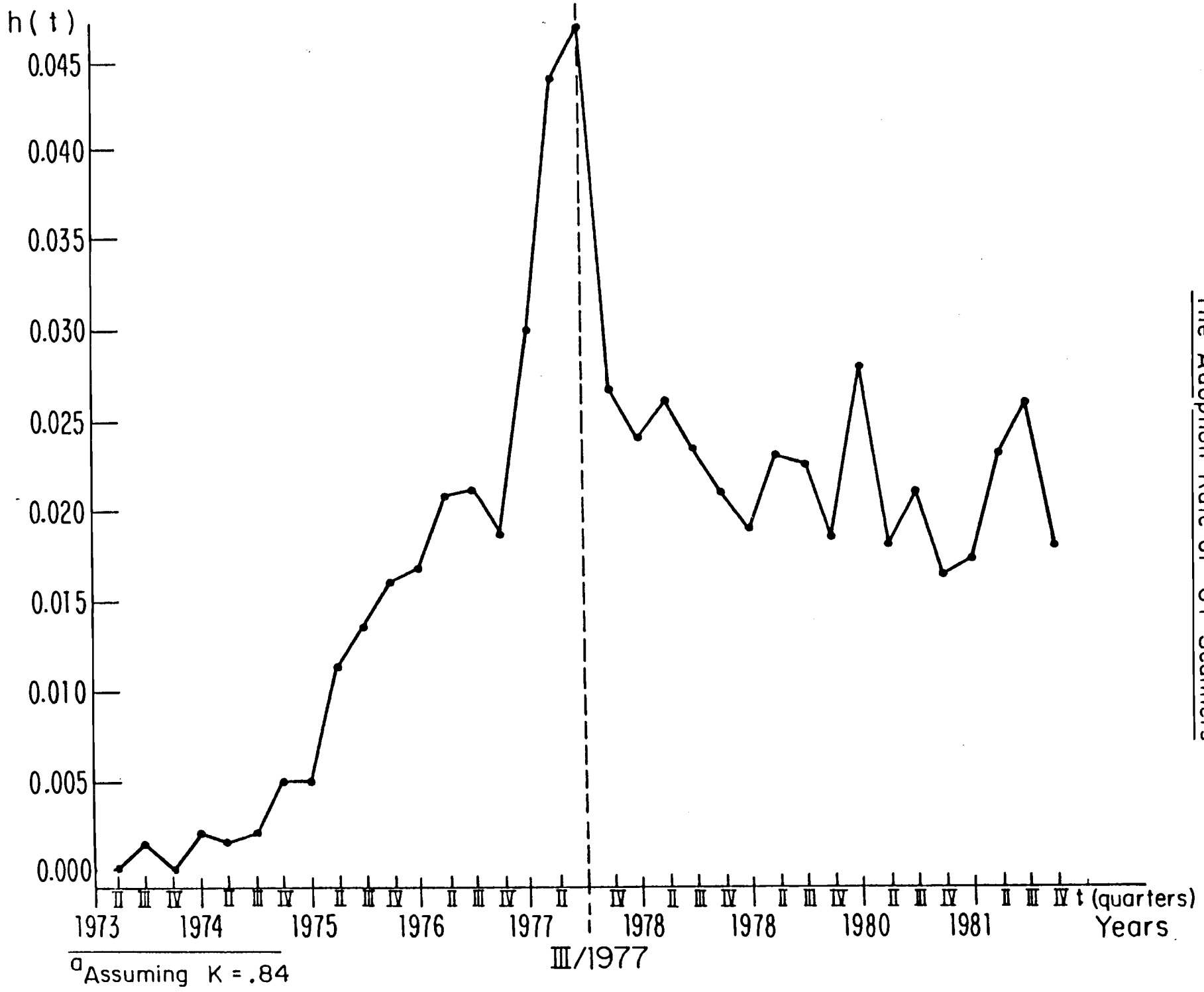
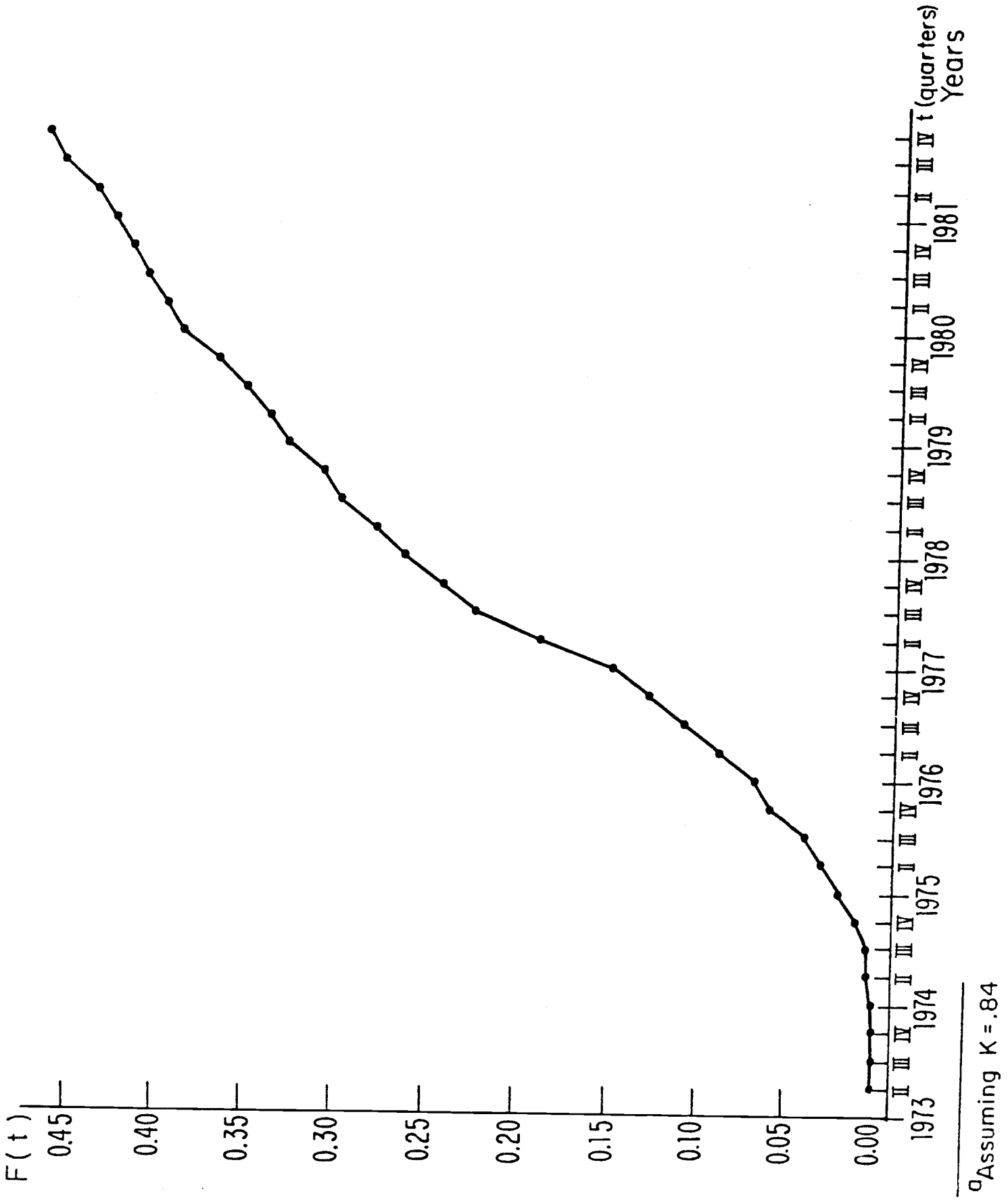



Figure 2
The Adoption Rate of CT Scanners[a]

[a]Assuming K = .84

Figure 3

The Diffusion Path of CT Scanners: F(t) [a]




[a] Assuming K = .84

Table VI

Logistic Estimates

| K | $\hat{\beta}$ | $r^2$ | $t^o = \frac{p/K}{\beta}$ (in years) |
|---|---|---|---|
| .84 | 0.059 | .85 | .657 |
| .70 | 0.061 | .86 | .760 |
| .60 | 0.063 | .88 | .860 |

7.3 Diffusion by Bed-Size Groups

As shown in Table II, the percentage of hospitals that adopted CT scanners by 12/81, $p_i$, varies a great deal across bed-size groups, ranging from .132 for hospitals with $100 \leq \text{beds} \leq 199$ to .876 for the largest hospitals. The question is how to compare them in terms of diffusion speed, in view of these disparities in the proportion of the process that is observed. It was argued in Section IV that there are two alternative ways to proceed: (a) truncate the processes at the same cut-off level (minimum $p_i$), or (b) make assumptions regarding $t_i^u$. In this case (a) has to be discarded because it will imply doing away with most of the available data. On the other hand we do not have specific priors regarding the unobserved segment of each group's process, and hence the only alternative left is to resort to the uniformity assumption, i.e., to estimate equation (22) for each group. This is certainly troublesome in the case of small hospitals in view of their very low $p_i$, and hence the estimates for them should be regarded accordingly.

Table VII presents both the estimates $\hat{r}_i$ and their components so as to provide a better idea of the nature of these calculations. Except for hospitals with 200-299 beds versus those with 100-199 beds, diffusion speed increases substantially with bed size, a result consistent with previous studies (e.g., Russell [20]). But, a closer look at the table indicates that the intra-group behavior of adopters was fairly similar across groups (as evidenced by the small variation in $r^*$ ), and that the observed differences in speed were due mostly to differences in $p/K$ and in the mean adoption time. To illustrate the point: suppose that the total number of hospitals with $100 \leq$ beds $\leq 199$ was much smaller, so that $p/K$ was equal to that of the largest hospitals; in that case diffusion in small hospitals would have been faster than in the 500+ group, because of the high $\bar{t}$ (i.e., small $(T - \bar{t})$ ) of the former, i.e., because the process was "crammed" in the later period. This underscores the fact that diffusion speed as defined here and elsewhere is only a measure of

Table VII

Estimates of Diffusion Speed under the Uniformity Assumption

| Bed-size | $p/K$[a] | $r^*$ | $(T - \bar{t})$ | $\hat{r}$ (months) | $\hat{r}$ (years) |
|---|---|---|---|---|---|
| 100 - 199 | .189 | 11.089 | 30.872 | 27.14 | 2.26 |
| 200 - 299 | .430 | 12.346 | 39.839 | 28.01 | 2.33 |
| 300 - 399 | .676 | 11.982 | 42.732 | 21.96 | 1.83 |
| 400 - 499 | .841 | 11.866 | 46.891 | 17.44 | 1.45 |
| 500+ | .894 | 11.476 | 57.527 | 16.36 | 1.36 |

[a]The p's are taken from Table II and the K's from Table III.

dispersion, and by no means the only relevant aspect of diffusion: the location in time of the process is at least as important, for which the mean adoption time, $\bar{t}$, is in fact an appropriate measure.[30]

## 7.4 Decomposition by Type of Control

The objective here is to illustrate the procedure developed in Section 5 for assessing the contribution of different groups to the Gini of the overall process. We have chosen for that purpose to partition the hospitals by type of control, as shown in Table VIII.

In view of the fact that there are no wide differences in the $p_i$'s, it is appropriate in this case to truncate the processes at $\min_i p_i = .3$,

Table VIII

Distribution of Hospitals[a] by Type of Control

| Type of Control | Number of adopters | Number of Hospitals | $p_i$ |
|---|---|---|---|
| Investor-owned/for profit (FP) | 102 | 342 | .30 |
| Government, non-federal (GNF) | 188 | 576 | .33 |
| Not-for-profit (NFP) | 912 | 2,160 | .42 |
| TOTAL | 1,202 | 3,078 | .39 |

[a]Community hospitals with beds > 100 .

30. The "origin" in Griliches [10], defined as $(-2.2 - \alpha)/\beta$, is precisely the mean adoption time as estimated by the logistic $(-\alpha/\beta)$, plus an arbitrary constant.

computing the $\Gamma^*_i$'s for the .3x (number of hospitals in class i ) initial observations of each group. The resulting figures are presented in Table IX and the decomposition in Table X (see equations (25) through (28)).

Table IX

Diffusion Parameters by Type of Control,

With Common Cut-off Level (p = .3)

| Control | (1) $n_i$ | (2) $n_i/N$ | (3) $\Gamma^*_i$ | (4) $T_i$ | (5) $\bar{t}_i$ |
|---|---|---|---|---|---|
| FP | 102 | .11 | 12.07 | 108 | 68.03 |
| GNF | 173 | .19 | 11.12 | 98 | 58.35 |
| NFP | 648 | .70 | 8.14 | 78 | 51.62 |
| All hospitals | 923 | 1.00 | 9.55[a] | 108 | 54.69 |

[a]This figure corresponds to $\Gamma_o$ in equation (24).

Table X

Decomposition of the Gini by Type of Control

| Control | (1) $\Gamma_{o.i}$ | (2) $w_i$ |
|---|---|---|
| FP | 3.843 | .40 |
| GNF | 3.322 | .35 |
| NFP | 2.384 | .25 |
| Total | 9.549 | 1.00 |

The findings can be summarized as follows: Not-for-profit hospitals were the fastest to adopt, followed by government non-federal and investor-owned (Table IX, column (3)). In spite of NFP being the largest group (70%), its diffusion process was the least correlated with the overall process, and the opposite is true for FP hospitals (Table X, column (1)). As to the net effect on aggregate diffusion, NFP hospitals accounted for 25% of the overall Gini, GNF for 35% and FP for 40% (Table X, column (2)).

## 8. Extensions

One of the main methodological advantages of studying diffusion with the aid of tools that are well known and widely applied (rather than issue-specific), is that the analysis of the phenomenon can be readily extended by drawing from the literature in which these and related tools play a key role. The Gini is certainly a tool of that nature, and it opens up numerous possibilities for further research. We would like to suggest two such extensions: the first links diffusion with stochastic dominance, and the second with rank-order tests of linear hypotheses. Both topics require careful elaboration, but that lies beyond the scope of this paper; thus, we shall limit ourselves here to introducing the issues.

Although the treatment of diffusion has been carried out almost exclusively in positive terms, there is certainly a normative aspect to it namely, the assessment of the relative desirability of alternative diffusion process, in view of the costs and benefits of delaying adoption, and society's time

preferences.[31] The following highly simplified formulation of this issue suggests a line of inquiry that may prove fruitful.

Assume that the net social loss due to postponing the adoption of the innovation until time $t$ is $v(t)$, exhibiting the properties: $v(0) = 0$, $v' > 0$ and $v'' < 0$, and that the objective is to minimize the expected value of the loss, $E[v(t)]$, for all $t \in (0,\infty)$. This is in fact analogous to the problem dealt with in the literature on stochastic dominance: As shown in Yitzhaki [27], given two diffusion processes $F_1(t)$ and $F_2(t)$, the necessary conditions for second order stochastic dominance are: $\mu_1 \leq \mu_2$ and $\mu_1 - \Gamma_1 \leq \mu_2 - \Gamma_2$, where $\mu_i$ is the expected adoption time and $\Gamma_i$ the Gini of $F_i$, $i = 1,2$. Sufficient conditions can also be derived for distributions that intersect at most once, with the aid of the extended Gini. If the factors affecting the value of these parameters were known (e.g., the effect of tax incentives, regulation, market structure, etc.), then it could be possible to design optimal or second best diffusion policies.

The second extension has to do with an issue of prime concern in diffusion studies: the identification of the variables that affect the decision to adopt, and the timing of adoption. This has been approached in various ways in the literature: linear probability models (Russell [21]), discrete choice models (Sommers [25]), simple regressions with the estimated $\beta_i$'s as the dependent variable (Griliches [10], Mansfield [16]), etc. Consider now the following linear model:

---

31. The most serious difficulties of this kind of welfare analysis are the identification of those costs and benefits, and the modelling of the dynamic interactions between the expectations and consequent decisions of individual adopters, and the evolution (pricewise and technological) of the innovations over time.

$$(32) \qquad t_i = \sum_{j=1}^{J} \alpha_j x_{ij} + \varepsilon_i \quad , \quad i = 1, \ldots, N$$

where $t_i$ is the time of the $i^{th}$ adoption, $x_j$ , $j = 1, \ldots, J$ are the variables presumed to affect the adoption decision (e.g., the characteristics of the individual adopters or of groups of adopters, time-dependent attributes of the innovation and of the environment, etc.), and $\varepsilon_i$ is the error term assumed to be i.id. but not necessarily normally distributed. Substituting (32) for $t$ in (15),

$$(33) \qquad \Gamma = 2 \text{ Cov}[t, F(t)] = 2 \sum_{j=1}^{J} \alpha_j \text{ Cov}[x_j, F(t)] + 2 \text{ Cov}[\varepsilon , F(t)]$$

where $2\alpha_j \text{ Cov}[x_j, F(t)]$ is the contribution of the $j^{th}$ variable to the Gini, and $2 \text{ Cov}[\varepsilon , F(t)]$ is its unexplained portion. If instead of estimating (32) we substitute $\bar{F}_i = \frac{R_i}{n}$ for $t_i$ (where $R_i$ is the rank of $t_i$ ), we obtain in fact Bennett's model [4] for non-parametric tests of linear hypotheses.[32] I.e., this specification allows to perform $\chi^2$ tests of the null hypothesis Ho: $\hat{\alpha}_1 = \ldots = \hat{\alpha}_j = \ldots = \hat{\alpha}_L = 0$ , $L < J$ , which meaning can be best understood in the context of (33).

Although the relative merits of this approach vis a vis those mentioned above are yet to be examined, it is worth noting some of the features that make it attractive: it provides with a <u>direct</u> way for assessing the effect of exogenous variables on the diffusion process (rather than indirect or two-stage procedures as in Griliches and Mansfield), it does not require restrictive assumptions regarding the distribution of the error term, and it enhances the coherence and scope of the methodology presented here for the study of diffusion.

---

[32] See also Shirley [24] for a description of the model.

# Appendix

## The Gini of the Gompertz Distribution

Let $t$ be a random variable having a Gompertz distribution:

$$F(t) = a^{b^t} \tag{A1}$$

Inverting it,

$$t(F) = (1/\ln b)\ \ln (\ln F/\ln a) \tag{A2}$$

Applying equation (15):

$$\Gamma = \frac{2}{\ln b} \int_0^1 \ln (\ln F/\ln a)(F - 0.5)\ dF = \tag{A3}$$

$$= \frac{2}{\ln b} \int_0^1 (\ln |\ln F| - \ln |\ln a|\ )(F - 0.5)\ dF \Rightarrow$$

$$\Gamma = \frac{2}{\ln b} \int_0^1 \ln |\ln F|\ (F - 0.5)\ dF \tag{A4}$$

Evaluating numerically the integral in (A4),

$$\int_0^1 \ln |\ln F|\ (F - 0.5)\ dF \cong 0.35 \tag{A5}$$

Thus,

$$\Gamma \cong 0.7/\ln b \tag{A6}$$

## References


1. Aitchison, J. and J.A.C. Brown: The Lognormal Distribution. Cambridge: Cambridge University Press, 1963.

2. Bain, A.D.: The Growth of Television Ownership in the United Kingdom since the War: A Lognormal Model. Cambridge: Cambridge University Press, 1964.

3. Barlow, R.E. and F. Proschan: Statistical Theory of Reliability and Life Testing. New York: Holt, Rinehart and Winston, Inc., 1975.

4. Bennett, B.M.: "Rank-Order Tests of Linear Hypotheses," Journal of the Royal Statistical Society, Series B, 30 (1968), 483-489.

5. Chandra, M. and N.D. Singpurwalla: "Relationships between some Notions which are Common to Reliability Theory and Economics," Mathematics of Operations Research, 6 (1981).

6. Coleman, J.S., E. Katz and H. Menzel: Medical Innovation: A Diffusion Study. Indianapolis: Bobs-Merrill, 1966.

7. Dixon, R.: "Hybrid Corn Revisited," Econometrica, 48 (1980), 1451-1461.

8. Feller, W.: "On the Logistic Law of Growth and its Empirical Verification in Biology," Acta Biotheoretica, 5 (1940), 51-65.

9. Feller, W.: An Introduction to Probability Theory and Its Applications, Vol. 2, 2nd edition. New York: John Wiley & Sons, Inc., 1971.

10. Griliches, Z.: "Hybrid Corn: An Exploration in the Economics of Technological Change," Econometrica, 27 (1957), 501-522.

11. Jarvis, L.S.: "Predicting the Diffusion of Improved Pastures in Uruguay," American Journal of Agricultural Economics, 63 (1981), 495-502.

12. Jensen, R.: "Innovation Adoption and Diffusion when there are Competing Innovations," Working Paper No. 71, Department of Economics and Mershon Center, Ohio State University, April 1981.

13. Jensen, R.: "Adoption and Diffusion of an Innovation of Uncertain Profitability," Journal of Economic Theory, 1982 (forthcoming).

14. Kendall, M. and A. Stuart: The Advanced Theory of Statistics, Vol. 1, 4th ed. London: Griffin, 1977.

15. Lomnicki, Z.A.: "The Standard Error of Gini's Mean Difference," Annals Math. Statist., 23 (1952), 635-637.

16. Mansfield, E.: Industrial Research and Technological Innovation. New York: Norton, 1968.

17. Nair, U.S.: "The Standard Error of Gini's Mean Difference," Biometrika, 28 (1936), 428-436.

18. Pakes, A.: "A Simple Aggregate Diffusion Model, Estimation Techniques, and a Case Study in a Secondary Country," Unpublished Ph.D. Dissertation, Chapter 4, Department of Economics, Harvard University, 1978.

19. Reinganum, J.F.: "Market Structure and the Diffusion of New Technology," The Bell Journal of Economics, 12 (1981), 618-624.

20. Russell, L.B.: "The Diffusion of Hospital Technologies: Some Econometric Evidence," The Journal of Human Resources, 12 (1977), 482-502.

21. Russell, L.B.: Technology in Hospitals: Medical Advances and their Diffusion. Washington, D.C.: The Brookings Institution, 1979.

22. Romeo, A.: "Interindustry and Interfirm Differences in the Rate of Diffusion of an Innovation," The Review of Economics and Statistics, 57 (1975), 311-319.

23. Shalit, H. and S. Yitzhaki: "Mean-Gini, Portfolio Theory and The Pricing of Risky Assets," Discussion Paper No. 8204, The Center for Agricultural Economic Research, Rehovot, Israel, 1982.

24. Shirley, E.A.: "A Distribution-free Method for Analysis of Covariance Based on Ranked Data," Appl. Statist., 30 (1981), 158-162.

25. Sommers, P. "The Adoption of Nuclear Power Generation," The Bell Journal of Economics, 12 (1981), 293-291.

26. Yitzhaki, S.: "Relative Deprivation and Economic Welfare," European Economic Review 17 (1982), 99-113.

27. Yitzhaki, S.: "Stochastic Dominance, Mean Variance and Gini's Mean Difference," American Economic Review, 72 (1982), 178-185.