Entry and Exit Echoes[*]

Boyan Jovanovic and Chung-Yi Tse[†]

July 29, 2009

Abstract

While aggregate data do not show the investment echoes predicted by vintage-capital models, echoes arise in rates of entry and exit of firms at the industry level. Moreover, industries where prices decline rapidly experience early 'shakeouts'. The relation emerges naturally in a vintage-capital model in which exit of firms sometimes accompanies the replacement of their capital, and in which a shakeout is the first replacement 'echo' of the capital created when the industry is born.

1   Introduction

In a class of models of embodied technological change, capital is periodically replaced.[1] In the same class of models, any burst of investment activity creates an "echo" when the investment becomes obsolete.[2] Faster technological progress then implies faster replacement of capital and, hence, more frequent investment echoes.

We argue that this by now standard vintage capital model has strong predictions for industry lifecycle dynamics — the existence of 'shakeouts' and 'echoes.' In particular, our model links these features to the rate of technological progress. We find that the data do support the 'vintage capital' view.

Two key assumptions deliver the simple link between technological progress and entry and exit echoes. First, we assume that the need to replace capital is associated

[†]NYU and Hong Kong University, respectively.

[1]Early examples are Johansen (1959), Arrow (1962) who model steady states.

[2]Boucekkine, Germain and Licandro (1997) and Mitra, Ray and Roy (1991) study dynamics in this class of models.
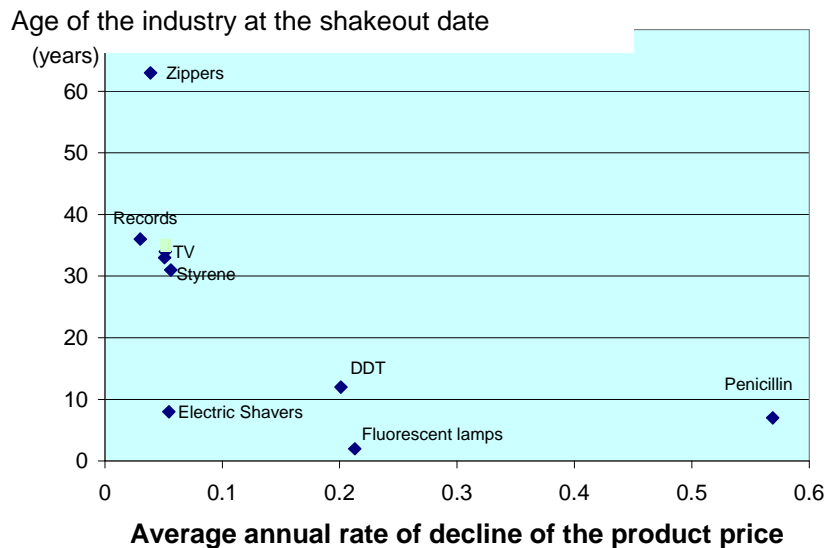
Figure 1: Technological progress and industry age at shakeout

with exit of firms faced with that need. We shall provide evidence from the airline industry that suggests this link is empirically important. Second, we assume that when an industry comes into being, a burst of investment is required. When that capital comes up for replacement, the investment echo then generates exit. We interpret this exit as the shakeout.

When an industry is competitive, technological progress leads to a reduction in the price of the output. Competitive producers, in other words, pass on the benefits of technological change to consumers. This means that we can proxy industry-specific technological change by the reduction in the product price. At first pass, the argument looks plausible: Figure 1 which shows that industries where prices decline rapidly experience early 'shakeouts' – simultaneous exits of a fraction of incumbent firms – just as the amended vintage-capital model implies.[3] This pattern will stand up when more data are added. Moreover, repeated echoes in entry and exit occur in several industries, and the interval of time between the first and second echoes is also shorter in industries where prices decline more rapidly: The faster the technological progress, the more frequent the echoes.

---

[3]The Figure uses data from Gort and Klepper (1982, henceforth GK). GK measure an industry's age from the date that the product was commercially introduced, i.e., from the date of its first sales. The shakeout period is defined as the epoch during which the number of firms is declining. GK say that an 'exit' occurs when a firm stops making the product in question, even if that firm continues to make other products. GK time the start of the shakeout when net entry becomes negative for an appreciable length of time. The shakeout era typically begins when the number of producers reaches a peak and ends when the number of producers again stabilizes at a lower level.

Ours is primarily a model of investment echoes. To make it a theory of exit echoes then requires explaining why capital replacement may sometimes cause exit. Our explanation is that as firms age, they sometimes lose the ability to implement new technology (which in turn is embodied in new capital). The assertion is roughly that 'old dogs can't learn new tricks'. We do not explain why they cannot, we instead parameterize this tendency in the form of a hazard rate of losing implementation skill exogenously and randomly.[4] A loss of implementation skill does not induce an exit right away. Rather, it induces an exit when all the capital that a firm owns reaches replacement age.

We also provide direct new evidence that firms with old capital are more likely to exit: In the air-transportation industry, exiting firms have capital that is on average eight years older than the capital of the surviving firms; in Section 5 we shall display this highly significant relation both for the U.S. and for the world as a whole. A related pattern emerges at the two-digit-industry level: Sectors that face more rapidly declining equipment-input prices experience higher rates of entry and exit (Samaniego, *forthcoming*). In other words, a sector that enjoys a high rate of embodied technological change will have a high rate of entry and exit, or what one would normally understand to be a higher level of creative destruction.

All technological change in our model is embodied in capital, which means that TFP should be constant when one adjusts inputs for quality. During the shakeout, a fraction of the capital stock is replaced by new capital; the number of efficiency units of capital stays the same but the productivity of capital per physical unit rises.

The model is of a standard vintage-capital type; Mitchell (2002) and Aizcorbe and Kortum (2005) use it to analyze industry equilibrium in steady state, i.e., the stationary case in which the effect of initial conditions has worn off, and after any possible investment spikes that may be caused by initial conditions have vanished. Jovanovic and Lach (1989) use it to analyze transitional dynamics but they generate neither a shakeout nor repeated investment echoes. We shall derive damped investment echoes that relate to the constant investment echoes derived by Boucekkine, Germain and Licandro (1997, henceforth BGL) in a similar GE model and by Mitra, Ray and Roy (1991) in a model where there is no progress but in which capital is replaced because it wears out. A number of other papers relate Figure 1, among them, Caballero and Hammour (1994), Klepper (1996), and Utterback and Suarez (1993). Some relate also to the repeated echoes that we shall document below. We shall discuss this work in Section 6.

Plan of the paper.—Section 2 presents the model. Section 3 and 4 describe tests of the two main propositions. Section 5 links capital replacement to firm exit empirically. Section 6 discusses other models in some detail and section 7 concludes the paper. The Appendix contains some proofs.

---

[4]Deeper reasons for why firms find it hard to adopt new technology are modeled by Klepper and Thompson (2009) and Chatterjee and Rossi-Hansberg (2008).

2   Model

Consider a small industry that takes as given the rate of interest, r, and the price of its capital. The product is homogeneous, and technology improves exogenously at the rate g. To use a technology of vintage t, however, a firm must buy capital of that vintage. The productivity of vintage-0 capital is normalized to 1, and so the productivity of vintage-t capital is e^{gt}.

Each firm is of measure zero and takes prices as given. Let p be industry price, q industry output, and $D(p, t)$ the demand curve at date t, assumed continuous in both arguments. Production of the good becomes feasible at $t = 0$.

The price of capital is unity for all t. Capital must be maintained at a cost of c per unit of time; c does not depend on the vintage of the capital, nor on time. Capital cannot be resold to firms outside the industry; it has a salvage value of zero. We assume that willingness to pay at small levels of q is high enough to guarantee that investment will be positive immediately.[5]

Implementation costs.—Let τ denote the age of a firm and let $\varepsilon(\tau)$ be that firm's cost of installing a unit of frontier capital. Since the cost of all capital is unity, the total cost of buying and installing capital of a τ-year old firm then is $1 + \varepsilon(\tau)$. In section 5, we will discuss evidence on how relative to their contribution to industry output, new firms implement new technologies more than incumbents do. This would happen if incumbents face some additional costs of adopting new technology. To model this tendency, we assume that $\varepsilon(\tau)$ is a Poisson process, independent over firms, with initial condition $\varepsilon(0) = 0$, Poisson parameter λ, and jump size $\kappa > 0$. This means that among firms aged τ, a fraction $e^{-\lambda \tau}$ will be on a par with entrants in their ability to adopt technology, and a fraction $1 - e^{-\lambda \tau}$ will face costs of at least $1 + \kappa$.

Industry output.—Capital is the only input. Let $K_t(s)$ be the date-t stock of active capital of vintage s or older, not adjusted for quality. Industry output at time t is the sum of the outputs of all the active capital

$$q(t) = \int_0^t e^{gs} dK_t(s). \tag{1}$$

Capital ceases to be active when it is scrapped.

Evolution of the capital stock.—Because capital is supplied to the industry at a constant price, the investment rate will exhibit damped echoes. Any mass point that occurs will repeat itself, though in a damped form. That is, if a mass-point of investment ever forms, will recur at a periodicity of T, and the size of the mass point will diminish over time. Moreover, there must be an initial mass point at $t = 0$ because no capital is in place when the industry comes into being. After that initial mass point, capital evolves smoothly until date T when the original capital

---

[5]Sufficient for this is that $\int_0^\infty e^{-rt} D^{-1}(0, t)\, dt > 1$, where $D^{-1}(q, t)$ denotes the inverse demand curve.
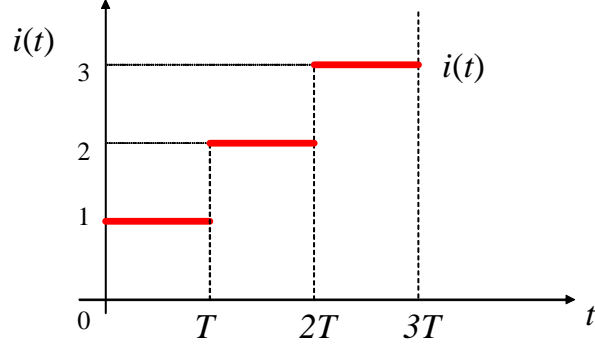
Figure 2: THE NUMBER OF SPIKES, $i(t)$.

is completely replaced by vintage-T capital. This is the second industry investment spike. The third investment spike then occurs at date $2T$, when all the vintage-T capital is replaced, and so on. Since the inter-spike waiting times are T, and since the first spike occurs at $t = 0$, the date of the i'th investment spike is $(i-1)T$, for $i = 1, 2, ...$ Let $i(t)$ be the integer index of the most recent spike.[6] We plot $i(t)$ in Figure 2.

We shall show that equilibrium is indeed of the form described in the previous paragraph: All the capital created at one spike date is replaced at the following spike date, T periods later. Therefore the capital stock at date t comprises capital created at the most recent spike date $i(t)$, plus the flow of investment, $x(t)$, over the preceding T periods.

Let X_i denote the size of the i'th investment spike. At date t, then, the amount of capital accounted for by the last spike is $X_{i(t)}$, and the date-t *cumulative* distribution of capital by vintage is

$$K_t(s) = X_{i(t)} + \int_{\max(0, t-T)}^{s} x(u)\, du. \tag{2}$$

We portray $K_t(s)$ in Figure 3. It has exactly one discontinuity at $(i(t) - 1)T$.

2.1 Equilibrium

The definition of equilibrium is simple if $x(t) > 0$ for all t. BGL call this the 'no holes' assumption because when it holds, the vintage distribution of capital in use has no gaps in it. To ensure a positive investment flow at all dates, i.e., $x(t) > 0$ all t, it is necessary and sufficient that output is increasing at all $t > 0$:

$$\frac{d}{dt} D(p, t) = \frac{\partial D}{\partial p} \frac{dp}{dt} + \frac{\partial D}{\partial t} > 0. \tag{3}$$

---

[6]Formally, $i(t) = \max\{i \in \text{ integers} \mid i \geq 1 \text{ and } (i-1)T \leq t\}$.
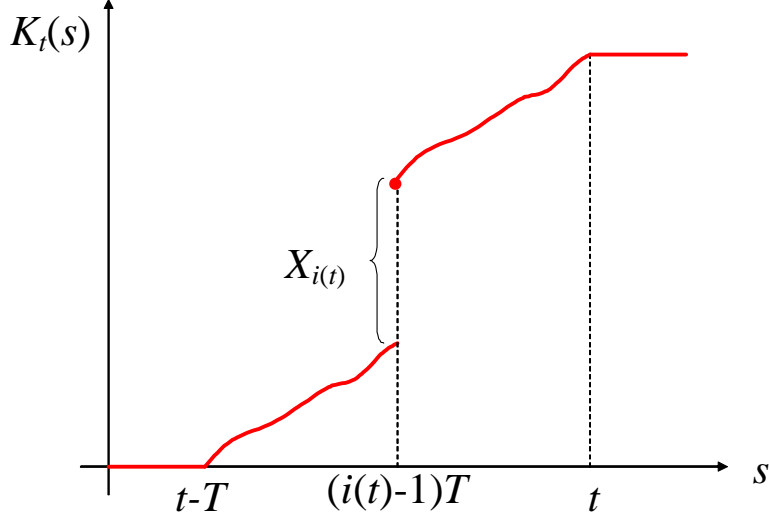
Figure 3: THE DATE-t DISTRIBUTION OF CAPITAL BY VINTAGE, s.

Thus it is possible for the time-derivative of D to be negative, but not too negative to offset the positive effect on output of the fall in price. Now, condition (3) is of limited value because it involves the endogenous variable p. One can, however, reduce (3) to a condition on primitives in some special cases. Take the case where population grows at the rate γ_t and where each consumer's demand is iso-elastic, i.e., the case where $D(p,t) = Ap^{-\lambda} \exp\left(\int_0^t \gamma_s ds\right)$. In an equilibrium in which price falls at a constant rate g, $dp/dt = -gp$. Then (3) is equivalent to

$$g\lambda > -\gamma_t$$

for all $t \geq 0$. Thus (3) can hold even if population is declining, as long as its rate of decline, γ_t, never exceeds $g\lambda$, or that demand be sufficiently elastic to offset any exogenous demand declines.

DEFINITION: *A constant-T equilibrium* consists of a product-price function $p(t)$, a retirement-age of capital, T, investment flows $x(t) > 0$, and investment spikes X_i accruing at dates $(i-1)T$, $(i = 1, 2, ...)$ that satisfy, for each $t \geq 0$, the following three conditions:

1. *Optimal retirement of capital*: The revenue of a vintage-t machine at date $t' \in [t, t+T]$ is $e^{gt}p(t')$. Since price declines monotonically, it is optimal to replace vintage-t capital as soon as its revenue equals its maintenance cost:

$$e^{gt}p(t+T) = c. \tag{4}$$

2. *Optimal investment*: Only entrants and incumbents for whom $\varepsilon(\tau) = 0$ will invest, because for them the total cost of a new machine is unity. If investment

6

$x(t) > 0$, the present value of a new capital good must equal its cost. The present value of the net revenues derived from that (vintage-t) unit of capital must satisfy

$$1 = \int_t^{t+T} e^{-r(s-t)} \left(e^{gt} p(s) - c \right) ds. \tag{5}$$

If the RHS of (5) were ever less than unity, $x(t)$ would be zero.[7]

3. *Market clearing*:

$$D(p(t), t) = e^{g(i(t)-1)T} X_{i(t)} + \int_{\max(0, t-T)}^t e^{gs} x(s) ds. \tag{6}$$

Since the industry does not exist before date zero, $x(t) = 0$ for $t < 0$, so that $D(p(0), 0) = X_1$.

Proposition 1 *If demand satisfies (3), a constant-T equilibrium exists, with*

$$p(t) = p_0 e^{-gt}, \tag{7}$$

where

$$p_0 = ce^{gT}. \tag{8}$$

and where T uniquely solves

$$\left(\frac{r+c}{c} \right) (r+g) = ge^{-rT} + re^{gT}. \tag{9}$$

Proof. (i) First we show that (7), (8), and (9) imply (4) and (5): Eq. (7) and (8) imply (4) for all $t \geq 0$, i.e., exit occurs at all $t \geq T$. When substituted into (5), (7) and (8) also imply $1 = c \int_t^{t+T} e^{-r(s-t)} \left(e^{g(t-s)} e^{gT} - 1 \right) ds$, which is equivalent to

$$1 + c\frac{1 - e^{-rT}}{r} = ce^{gT} \frac{1 - e^{-(r+g)T}}{r+g}. \tag{10}$$

Multiplying by $r(r+g)/c$ we have $r(r+g) + c(r+g)\left(1 - e^{-rT}\right) = cr\left(e^{gT} - e^{-rT}\right)$, and cancelling cre^{-rT} and combining terms we reach (9). Therefore if T solves (9), (5) holds.

(ii) Exactly one solution T to (9): The LHS of (9) is constant and exceeds $r+g$. On the other hand, the RHS of (9) is continuous and strictly increases from $r+g$ to infinity, having the derivative $rg\left(e^{gT} - e^{-rT}\right) > 0$ for $T > 0$. Therefore the two curves have exactly one strictly positive intersection.

(iii) If (3) is satisfied with $dp/dt = -gp$, there exists a sequence of $x(t) > 0$ that satisfies (6). ∎

---

[7]Proposition 4 of BGL covers that case which arises when there is too high an initial stock of capital. This cannot be true at $t = 0$ in our model, and we shall state conditions in (3) that exclude it as an equilibrium possibility at any date.

2.2 The nature and frequency of the spikes

Relation to Figure 1.—The negative pattern arises across steady states indexed by g; a situation in which each industry experiences its specific rate of technological progress g. Implicitly differentiating (9) in the Appendix, we reach the result:

Proposition 2 *The replacement cycle is shorter if technological progress is faster:*

$$\frac{\partial T}{\partial g} = -\frac{\frac{r+c}{c} + (gT - 1) e^{gT}}{g^2 \left(e^{gT} - e^{-rT}\right)} < 0. \tag{11}$$

Thus, industries with higher productivity growth should have an earlier shakeout and more frequent subsequent spikes.[8]

The second relation concerns the rate at which investment echoes or investment spikes die off:

Proposition 3 *Investment spikes decay geometrically. That is,*

$$X_n = e^{-gT(n-1)} X_1 \tag{12}$$

for $n = 1, 2, ...$

Proof. Because $p(t)$ is continuous at T, the number of efficiency units replaced at the spikes is a constant, X_1, which means that $X_n = e^{-gT} X_{n-1}$, and (12) follows. ∎

The spike dates remain T periods apart, at $0, T, 2T, 3T, ...$ The spikes occur regularly because technological progress occurs at the steady rate g. The assumption that g is constant is essential for both the constancy of T and for the geometric nature of the decay in the mass of investment. Asymptotically, however, the spikes vanish and the equilibrium becomes like the one that Mitchell (2002) and Aizcorbe and Kortum (2005) analyze.

Firm exit at the investment spike.—At a replacement spike at date t, say, all firms that need to replace their capital bought that capital at date $t - T$. All firms, whatever their age was at date $t - T$ had $\varepsilon = 0$ at that time Among those, a fraction $1 - e^{-\lambda T}$ will have experienced a jump in ε and will not find it profitable to invest. These firms will then exit.

The remaining parameters of the model are the maintenance cost c and the rate of interest r. The following claim is proved in the Appendix:

---

[8] We interpret g as an invariant property of an industry. But it can be interpreted as applying only to a given epoch in the lifetime of a given industry, and subject to occasional shifts. In Aizcorbe and Kortum (2005), e.g., one can think of such shifts as tracing out the relation between technological change and the lifetime of computer chips.

Proposition 4

$$(i) \quad \frac{\partial T}{\partial c} < 0. \quad And, \ if \ gc < r^2, \quad (ii) \quad \frac{\partial T}{\partial r} > 0, \quad (iii) \quad \frac{\partial (gT)}{\partial g} > 0.$$

A rise in the maintenance cost, c, reduces the lifetime of capital as one would expect. Since replacing capital constitutes an investment, when the rate of interest rises that form of investment is discouraged, and it will occur less frequently.

The role of the demand curve.—As long as the demand curve is sufficiently elastic and as long as exogenous forces do not lead demand to fall too much (as specified in (3)), only X_1 – industry capacity created at the outset, and the rate of investment to follow $x(t)$ depend on the demand for the product. Neither the time-path of prices, nor the frequency of the spikes depends on it. These results also rest on the assumption that capital is supplied perfectly elastically at a unit price, an assumption that seems reasonable if the industry at hand is a small fraction of the economy or if, as in a one-sector growth model, capital is produced linearly, at constant returns to scale.

3 T vs. g: Testing Proposition 2

We shall begin by proxying the age, T, of an industry at its first replacement spike, by the industry's age at which the shakeout of its firms begins. This measure will underlie the first test of the model. And as suggested by (7), we shall proxy the rate of technological progress, g, by the rate at which the price of the product declines. Since replacement episodes are in the model caused by technological progress, we first check if industries with higher productivity growth experience earlier shakeouts. That is, we ask whether industries with a high g have a low T as Proposition 2 claims and, if so, how well the solution for T to (9) fits the cross-industry data on g and T.

We now describe the procedure by which we choose the model's parameters. By Proposition 1, a unique solution to (9) for T exists; denote it by $\tilde{T}(r, c, g)$. In Table 1, $\hat{T}^{(1)}$ is the date that GK find that Stage 4 (the shakeout stage) begins in their various industries. In a couple of cases the shakeout had not yet begun, and they are censored. The annual rate at which the price declines, averaged over the period $\left[0, \hat{T}^{(1)}\right]$ is $\hat{g}_i^{(1)}$. We do not have observations on r and c; we find $r = .07$ a reasonable guess, but wish to check robustness with respect to this assumption and therefore choose three alternative values for r, namely $.02, .07$ and $.12$ and in each case estimate c. The estimation assumes that the distribution of c over industries is log-normal: For all firms in industry i, $\ln c_i$ is a draw from $N(\ln \bar{c}, \sigma^2)$. We then estimate $\bar{c}$ and σ to fit the model to the data; the details are in the Appendix.

The data and the three implicit functions $\tilde{T}(g, r, c)$ fitted to them are shown in Figure 4.[9] The estimates of $\bar{c}$ are reported in the figure. Since r makes little difference

---

[9] Table 7 of GK reports the number of years until shakeout for 39 industries. But only for 8 of

to the predicted relation we shall proceed with the value $r = .07$ for the remainder of the analysis.

PRODUCT NAME	$\hat{g}^{(1)}$	$\hat{T}^{(1)}$	$\hat{g}^{(2)}$	$\hat{T}^{(2)}$
Auto tires			0.03	21
Ball-point pens	0.07	>28	0.07	>18
CRT			0.04	24
Computers			0.07	19
DDT	0.20	12	0.13	16
Electric shavers	0.05	8	0.01	5.5
Fluorescent lamps	0.21	2	0.21	2.5
Nylon	0.03	>34	0.03	>21
Home & farm freezers			0.02	18
Penicillin	0.57	7	0.57	8.5
Phonograph records	0.03	36	0.02	8.5
Streptomycin			0.31	22
Styrene	0.06	31	0.07	12
Television	0.05	33	0.08	8
Transistors			0.17	15
Zippers	0.04	63	0.05	38

Table 1: THE DATA IN FIGURES 4 AND 5.[10]

Our estimate of c is in the range of typical maintenance spending. McGrattan and Schmitz (1999) report that in Canada, total maintenance and repair expenditures have averaged 5.7 percent of GDP over the period from 1981 to 1993, and 6.1 percent if one goes back to 1961. These estimates are relative to output, however, whereas ours are relative to the purchase price of the machine which is normalized to unity. Relative to output, maintenance costs are one hundred percent at the point when the machine is retired (this is equation [4]). Maintenance costs are constant over the machine's lifetime, whereas the value of the machine's output relative to the numeraire good is e^{gT} when the machine is new. For instance if $g = 0.1$ and $c = 0.05$, our model gives a value of $T = 16$; then $e^{gT} = 4.9$, so that as a percentage of output, maintenance spending ranges between 20 and 100 percent. Therefore, c must stand partly for wages to workers as a fixed-proportion input as in the original vintage-capital models like Arrow (1962) and Johansen (1959) that had a fixed labor requirement.

them does Table 5 report data on the rate of price decline up to the shakeout. These are the eight reported in Figures 1 and 4.

[10]The variable definitions are: $\hat{T}^{(1)}$ = age at the start of shakeout, $\hat{g}^{(1)}$ = average annual price decline on $\left[0, \hat{T}^{(1)}\right]$, $\hat{T}^{(2)}$ = the number of years elapsed between the start of stage 2 and the midpoint of the shakeout stage, and $\hat{g}^{(2)}$ = annual price decline over stages 2 and 3.
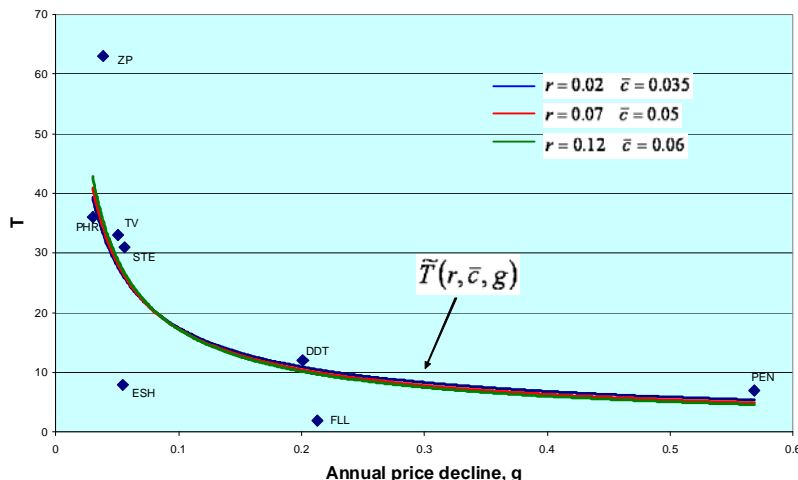
Figure 4: THE RELATION BETWEEN $\hat{T}$ AND $\hat{g}$ FOR THE EIGHT UNCENSORED GK OBSERVATIONS

3.0.1 Testing Proposition 2 using an alternative definition of T

We now entertain a different definition of T; one that may provide a better test of the model, and one that will provide us with more observations. The main reason for doing this is that GK's 'Stage 1,' defined as the period during which the number of producers is still small (usually two or three), may not contain what we would call an investment spike. During stage 1, only a few firms enter, a number that is in some industries – autos and tires, e.g., – much smaller than the number of firms that exit during the shakeout.

The model predicts a date-zero investment spike $X_1 = D(p_0, 0)$, without which there would be no exit spike at date T. Not all the GK industries will fit this, however. Indeed, GK state that rarely is a product's initial commercial introduction immediately followed by rapid entry. Autos, e.g., had very low sales early on, and it took years for sales to develop.[11] Therefore the spike is better defined at or around the time when the entry of firms was at its highest. Moreover, in GK, for many industries, the shakeouts were not completed until a few years after they first began. It may thus be more appropriate to designate the shakeout date as the midpoint of the shakeout episode instead of as the start date of the shakeout episode.

In light of this, let $\hat{T}^{(2)}$ be the time elapsed between the industry's 'Takeoff' date (which is when Stage 2 begins) and the midpoint of the the industry's shakeout episode (the midpoint of Stage 4). This revised definition for $\widehat{T}$ calls for adjusting

---

[11]Klepper and Simons (2005), however, do find an initial spike for TVs and Penicillin – both start out strong after WW2. There may be a problem with the TV birthday being set at 1929 as GK have it. During WW2 the Government had banned the sale of TVs.
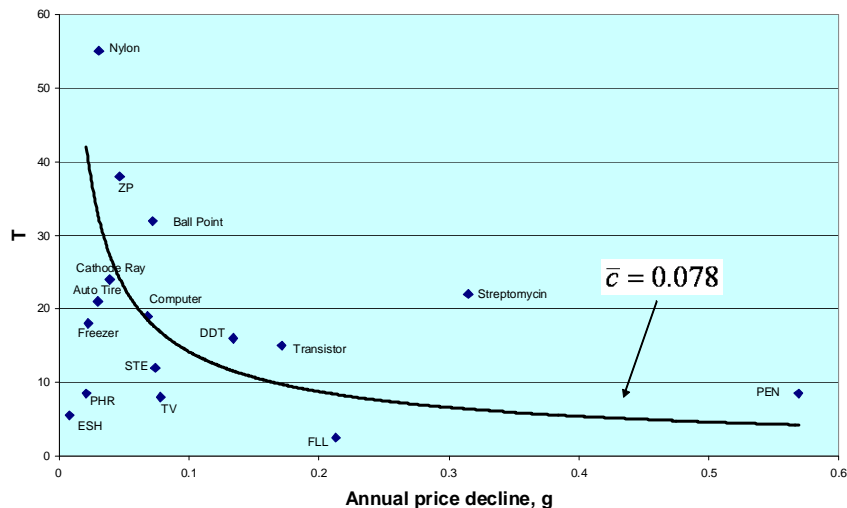
Figure 5: $\hat{T}$ AND $\hat{g}$ USING THE ALTERNATIVE DEFINITION OF T

$\hat{g}$ to be the rate of the average annual price decline between the takeoff date (which comes after the industry has completed stage 1), and the first shakeout date. We call this variable $\hat{g}^{(2)}$. This allows us to enlarge the sample. The additions are as follows:

1. For six of the GK industries, complete information for price-declines in Stages 2 and 3 (but not Stage 1) is available. They can now be added to the analysis.

2. The two censored observations listed in Table 2 will also be added.[12]

3. We replace the GK information for the TV by that reported in Wang (2006), who compiled the data from the Television Factbook. This change may better reflect the history of the TV industry as GK dated the birth of the industry as early as 1929, while according to Wang, the commercial introduction of TV starts only in 1947.

Figure 5 plots $\widetilde{T}$ for the various industries.[13] The estimate of $\bar{c} = 0.078$ is a bit higher than with the estimate in the previous sample ($\bar{c} = 0.05$). This is because T is smaller under the new definition, and a higher replacement cost is needed to generate the earlier replacement.

---

[12] The likelihood function is amended accordingly. See Appendix for details.

[13] For the two truncated observations, Ball–point pens and Nylon, the data points in Figure 5 represent the respective means conditional on their truncated values. The Appendix explains how the conditional means are computed.

4   Echoes: Testing Proposition 3

Let us first elaborate on the content of Proposition 3.

1. Shakeouts should diminish geometrically in absolute terms. Relative to the size of the industry, however, the decline will be faster than (12) indicates because industry output increases from its date-zero level – both because price declines and because the demand curve tends to shift to the right over time.

2. Moreover, exit spikes should coincide with entry spikes.

3. The inter-spike waiting times should depend negatively on g as Proposition 2 claims.

This section tests these implications using Agarwal's extension and update of the GK data, described fully in Agarwal (1998).[14] The products are listed and some statistics on them presented in Table 3 in the Appendix.

A procedure for detecting spikes must recognize the following features of the data:

(A) *Length of histories differ by product.*—Coverage differs widely over products, from 18 years (Video Cassette Recorders) to 84 years (Phonograph Records).

(B) *The volatility of entry and exit declines as products age.*—The model predicts that the volatility of entry and exit should decline with industry age. Other factors also imply such a decline: (i) Convex investment costs at the industry level, as in Caballero and Hammour (1994), and (ii) Firm-specific c's. Both (i) and (ii) would transform our X_n from spikes into waves and, eventually, ripples.[15]

Hodrick-Prescott residuals in entry and exit rates.—Our spike-detection procedure is in the spirit of the investment-spike literature that defines a spike as an unusually high investment rate.[16] Roughly speaking, we shall say that a spike in a series Y_t occurs whenever its HP residual is more than two standard deviations above its mean. 'Roughly', because of adjustments for (A) and (B) above. We constrain industry i's

---

[14]The evidence hitherto is mixed: Cooper and Haltiwanger (1996) find industry-wide retooling spikes, but GK did not report second shakeouts, though this may in part be because the GK data rarely cover industry age to the point $t = 2T$ where we ought to observe a second shakeout. In any case, Agarwal's data cover more years and contain entry and exit separately, and we shall use them to study this question.

[15]Spikes may also dissipate because (i) A positive shock to demand would start a new spike and series of echoes following it; these would mix with the echoes stemming from the initial investment spike, (ii) Random machine breakdowns at the rate δ would transform (12) into $X_n = e^{-(g+\delta)T(n-1)}X_1$, which decays faster with n.

[16]Gourio and Kashyap (2007) record a spike whenever investment exceeds twenty percent of the beginning-of-period capital stock, which is roughly 2.5 times the level of replacement investment. This leads to 15-20 percent of the years being spike years. Our procedure produces a weighted average of 4.3 percent of the years as exit-spike years and 5.6 percent as entry-spike years.

trend, τ, by

$$\sum_{t=2}^{A_i} (\tau_t - \tau_{t-1}) \leq a A_i^b,$$

where A_i is the age at which an industry's coverage ends. We set $a = 0.005$ for both series. Because both series are heteroskedastic, with higher variances in earlier years, we chose $b = 0.7$ for both entry and exit (If b were unity, an industry with longer coverage would have a larger fraction of its observations explained by the trend). The trend therefore explains about the same fraction of the variation in short-coverage industries as in long-coverage industries.

This fixes problem (A), but not (B): The HP residual, $u_t \equiv Y_t - \tau_t$, is still heteroskedastic, the variance being higher at lower ages. To fix this, we assume that the standard-deviation depends on product age as follows:

$$\sigma_t = \sigma_0 t^{-\gamma}, \tag{13}$$

where $\sigma_0 \geq 0$ and $\gamma \geq 0$ are product-specific parameters estimated by maximizing the normal likelihood[17]

$$\prod_{t=1}^{A_i} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{1}{2}\left[\frac{u_t}{\sigma_t}\right]^2\right).$$

The spike-detection algorithm.—If at some date the HP residual is more than two standard deviations above its mean of zero, then that date is a spike date. But we shall allow for the possibility that unusually high replacement will take up to three periods. Thus we shall say that a series Y_t in a certain time window is above 'normal' if one or more of the following events occurs

1-period spike:	$u_t > 2\sigma_t$,
2-period spike:	$u_t > \sigma_t$ and $u_{t+1} > \sigma_{t+1}$,
3-period spike:	$u_t > \frac{2}{3}\sigma_t$ and $u_{t+1} > \frac{2}{3}\sigma_{t+1}$ and $u_{t-1} > \frac{2}{3}\sigma_{t-1}$.

The cutoff levels of 2, 1, and $\frac{2}{3}$ times σ_t were chosen in the expectation that each of the three events would carry the same (small) probability of being true under the null. The latter depends on the distribution and the serial correlation of the u_t which we do not know. But, again for the normal case, these probabilities turned out to be roughly the same. That is,

$$1 - \Phi(2) = 0.023, \quad (1 - \Phi(1))^2 = 0.025, \quad \text{and} \quad \left(1 - \Phi\left(\frac{2}{3}\right)\right)^3 = 0.016.$$

[17]Although the HP residuals are not independent and unlikely to be normal, this procedure still appears to have removed the heteroskedasticity in the sense that the spikes were as likely to occur late in an industry's life as they were to occur early on.

Table 3 summarizes the results in more detail. The 33 products are listed alphabetically and are so numbered. To explain the table, let us focus on product 23, Phonograph Records and read across the row. Records were first commercialized, i.e., sold, in 1908. Being the oldest product, it also is the product for which we have the most observations, 84, since (with one exception) the series all end in 1991.[18] The next four entries are the exit and entry spikes, by age of industry and by calendar year. There are seven one-year spikes and one two-year spike, this being the last exit spike. The 1934 exit spike falls in the GK shakeout region the dates of which are in the last column of the table. See Figure 6 where the GK shakeout region is shaded. The remaining columns report the correlations between the entry and exit series. The raw series are negatively correlated – when the industry is young, entry is higher than exit, and later the reverse is true – but the correlation is slight (-0.07). The trends (i.e., the τ's) are more negatively correlated (-0.27). The HP residuals, on the other hand, are *positively* correlated; our model suggests that this should be so because the spikes should coincide. The correlations averaged across products are at the bottom of the table.

Figure 6 plots the exit and entry series for Phonograph records, in each case plotting the HP trend and the two-standard-deviation band. Spikes are circled. Let us note the following points:

1. For both entry and exit the spikes are evenly distributed as a function of industry age, predominating neither early nor late in the life of the industry. Table 3 shows this is true in most industries. This suggests that the heteroskedasticity adjustment in (13) is adequate.

2. The number of entry and exit spikes is equal – four entry and four exit spikes. But only the final, fourth spikes coincide in that they are within one year of each other. There were ten other products for which this was so. In seven out of the ten, there is at least one instance where an entry an exit spikes occur in the same year.

3. The second exit spike is well in the GK region, but there should also have been an entry spike in that region. Over all the industries the number of entry spikes (67 in all) is slightly less than the number of entry spikes (79 in all); for the nine industries in which the GK region does contain an Agarwal spike, it is *always* an exit spike – see the numbers in italics in Table 3.

4. Just as our model predicts, however, $T_{2,i} - T_{1,i}$ is negatively related to $\hat{g}_i$. That is, analogously to the result in Figures 4, the first exit spike is followed sooner by the second exit spike in those industries i where prices decline faster in the

---

[18]Since neither the entry nor the exit rate is defined in the year of the industry' birth, our entry and exit rate series begin in the year after the birth of the industry.

Phonograph Record Entry Rates

NBER = recession years

GK
Stage 4
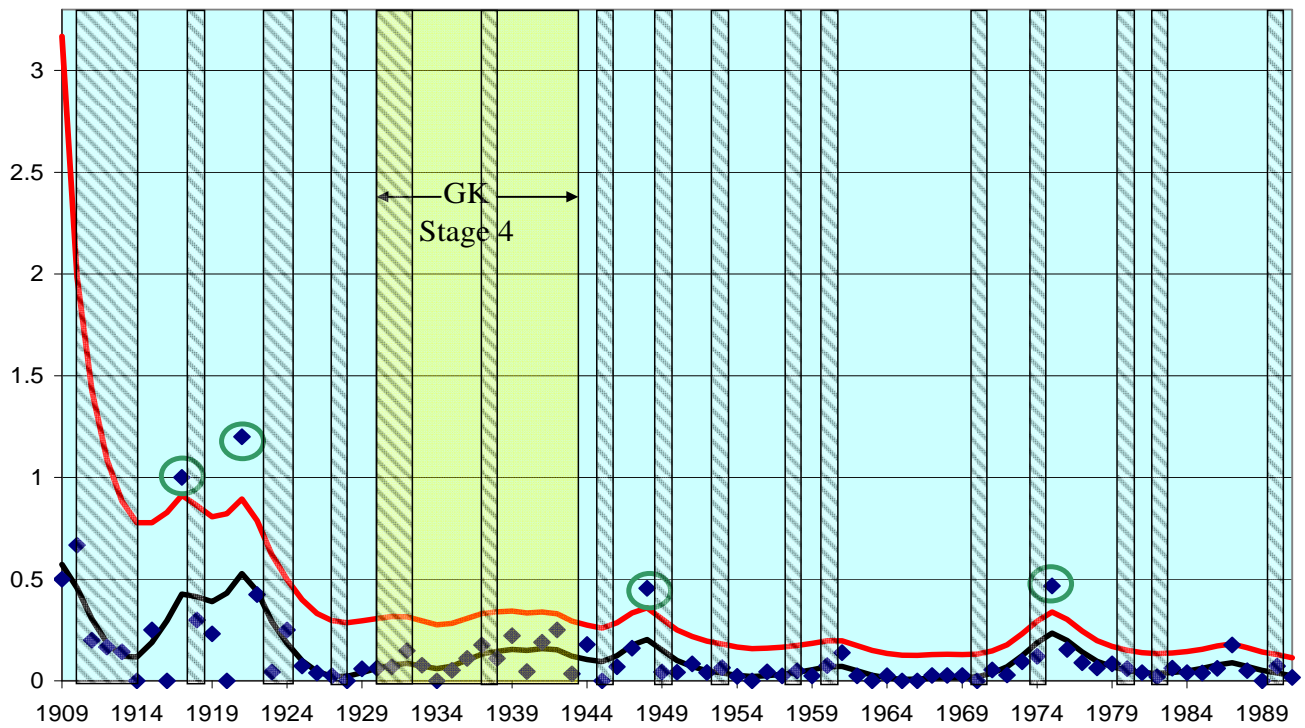
Phonograph Record Exit Rates

GK
Stage 4
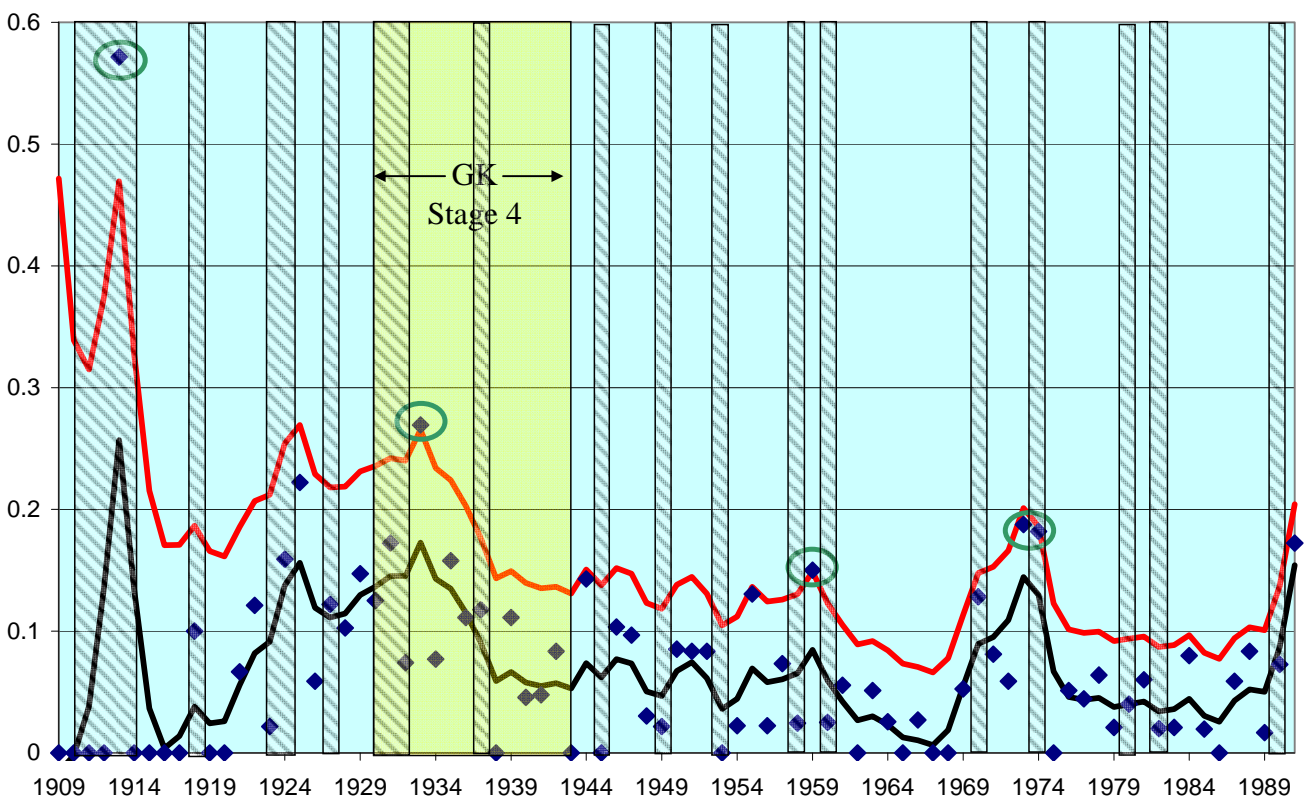
Figure 6: Spikes in the Phonograph-Record Industry
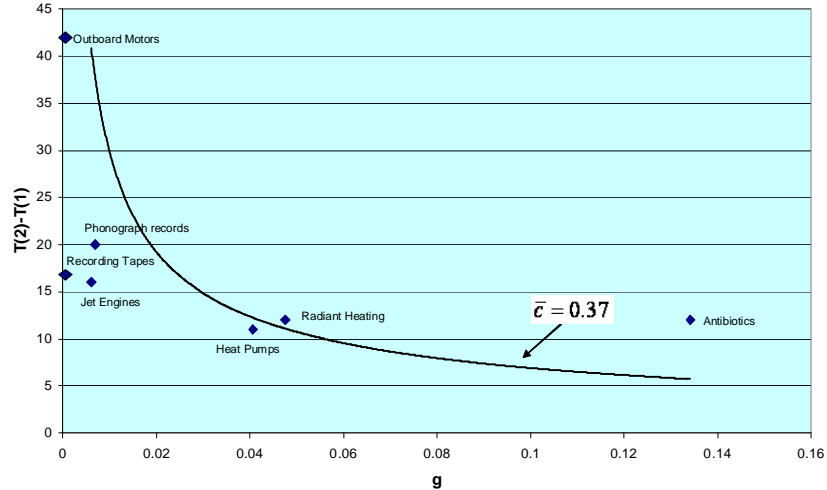
Figure 7: EXITS: THE RELATION BETWEEN $\hat{T}_2 - \hat{T}_1$ AND $\hat{g}$

interim. We now calculate $\widehat{g}$ as the rate of average price decline during the years covered by $T_2 - T_1$. To maximize the number of observations, we include products for which price information is available for as little as 70% of the time during the years covered by $T_2 - T_1$. For *Outboard Motors* and *Recording Tapes*, however, price went up during those years which is inconsistent with the model. But they both had above-average $T_2 - T_1$ values, 42 and 17 respectively, and while we did not include them in the estimation routine, we include them in Figure 7, setting $\hat{g}$ to zero in both cases.

5. The same test is also done on entry spikes. Once again, $T_{2,i} - T_{1,i}$ is negatively related to $\hat{g}_i$. The results are in Figure 8.[19]

4.0.2 Entry- and exit-spike coincidence

If the entry and exit spikes did not coincide, industry price would not fall at the rate needed to satisfy the free-entry condition. Thus entry and exit spikes in capital should occur simultaneously: Capital retired at the spike date should equal the capital brought in at that date. If the fraction of exiting capital accounted for by the exiting firms (as opposed to continuing incumbents) was the same as the fraction

---

[19]For both plots we assume, as before, that $\ln c_i \sim N\left(\ln \bar{c}, \sigma^2\right)$ in the estimation. Instead of fitting a $\widetilde{T}$ curve to the observations independently, a more rigorous procedure would perhaps be to calibrate a $\widetilde{T}$ curve with $\bar{c}$ set equal to the estimates in Figures 4 and 7. After all, the theory assumes a stationary c. But the products in the Agarwal dataset for which we can construct measures of $T_2 - T_1$ and g are not the same products in Figures 4 and 7 from the GK dataset. Our procedure for designating spike dates also differs from GK's procedure for designating shakeout dates.
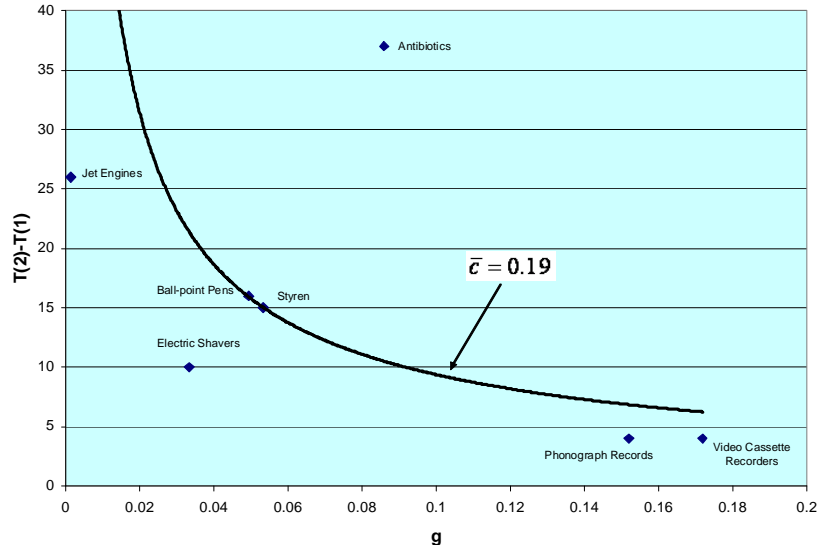
Figure 8: ENTRIES: THE RELATION BETWEEN $\hat{T}_2 - \hat{T}_1$ AND $\hat{g}$

of new capital accounted for by the new entrants (again, as opposed to continuing incumbents) then the capacity and output relinquished by the exiting firms should equal the capacity brought in by the entering capital. And, in particular, the market share of the exiting firms would equal the market share of the entrants. More generally, if the fractions were unequal but fixed over time, the model still predicts a positive correlation of entering and exiting capital over time and, hence a positive correlation between the market shares of entrants and exiters.

This implication gets support from Table 7 of Dunne, Roberts and Samuelson (1988, 'DRS') which reports that whether or not one controls for industry effects, the market shares of entering and exiting firms are positively correlated over time, when the unit of time is five-years. Because our data do not have information on market shares or on the output of individual entrants and exiters, we can deal only with firm numbers. For numbers of entrants and exits, Table 7 of DRS find that the numbers of entrants and exits is negatively correlated once they control for industry effects. Our data show a positive correlation, probably because they are much more finely disaggregated than the DRS data are.[20]

In our data, a period lasts one year, and entry and exit spikes coincide only in a minority of cases. We shall now test the null hypothesis that the entry and exit spikes are not correlated. We shall find that the number of coincidences highly significantly

---

[20]For instance, Antibiotics, the first product in Table 3, is just one of the 63 products classified under the 4-digit industry *Pharmacetical Preparations* (2834). And Electric Shavers, the 9th product in Table 3, is just one of 67 products classified under the 4-digit industry *Electric Housewares and Fans* (3634).

exceeds the fraction that would be expected to arise if the two sets of dates were uncorrelated.

Test of coincidence.—Under the null that the entry dates are uncorrelated with the exit dates, we now derive the probability that there will be no coincidences in dates. This will be a limited-information test because we are unable to derive the probability distribution of the full sample. Suppose an industry has τ periods of coverage and that during these τ periods there were N_E entry spikes and N_X exit spikes. Suppose entry spikes happened at dates $t_1, t_2, ..., t_{i_{\text{entry}}(\tau)}$. Suppose that the dates of the entry spikes were not correlated with the dates of the exit spikes. Suppose, moreover, that *an exit spike was equally likely to fall at any date.* No match will arise if and only if none of the N_X exit spikes matches any of the N_E entry spikes. In other words, under the null hypothesis, H_0, the spikes are exchangeable random variables which are correlated only in so far as they cannot fall at the same date as another spike. Exit-spike i is equally likely to fall in each 'free' year that the data cover. We condition on the entry-spike dates. Only their number matters, not their actual dates: Any set of dates would produce the same likelihood.

To derive the likelihood, begin by placing the exit spikes in random order $i = 1, ..., N_X$. Then under H_0: "entry and exit dates are uncorrelated", the probability that the first exit spike falls in one of the $\tau - N_E$ bins not occupied by an entry spike is $\frac{\tau - N_E}{\tau}$. Since no two exit spikes can occupy the same bin, conditional on this event, the probability that the second randomly chosen spike date will produce no match is $\frac{\tau - N_E - 1}{\tau - 1}$. Proceeding in this way sequentially until the N_X'th exit spike, the probability under H_0 that there is no match is

$$1 - \rho \equiv \prod_{i=1}^{N_X} \left(\frac{\tau - N_E - (i-1)}{\tau - (i-1)} \right). \tag{14}$$

Denote the solution for ρ by $\rho = \psi(\tau, N_X, N_E)$. The formula conditions on τ, N_X and N_E all three of which generally differ over industries. For industry j write

$$\rho_j = \psi\left(\tau^j, N_X^j, N_E^j\right). \tag{15}$$

Then ρ_j is the probability that industry j has at least one match. Now define the dichotomous random variable m_j as follows: Let

$$m_j = \begin{cases} 1 & \text{if industry } j \text{ has at least one exit-entry match} \\ 0 & \text{otherwise} \end{cases}$$

so that $m_j = 1$ with probability ρ_j and $m_j = 0$ with probability $1 - \rho_j$. That is, m_j is binomially distributed.
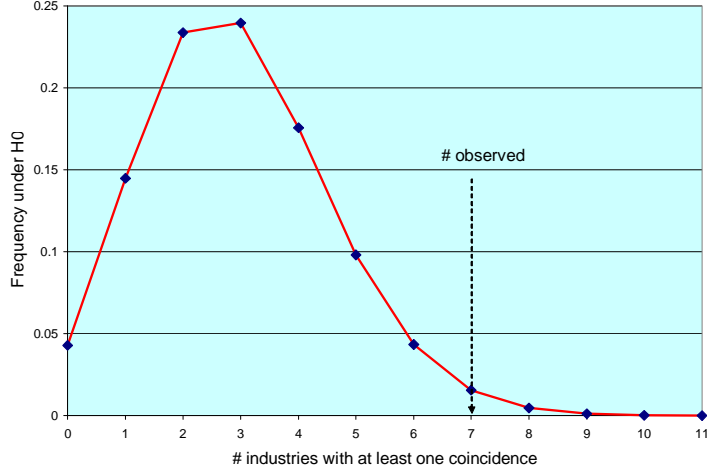
Figure 9: FREQUENCY DISTRIBUTION OF M UNDER H_0

The Likelihood function. – Let ρ_j be the parameter for industry j, defined in (15). Among the 31 industries for which ρ_j can be defined, let J be the set of industries for which $m_j = 1$. The likelihood of the sample is

$$\mathcal{L}\left(m_1, ..., m_{31}\right) = \prod_{j \in J} \rho_j \prod_{j \notin J} \left(1 - \rho_j\right) = 4.6852846 \times 10^{-9}. \tag{16}$$

The p-value.—The p-value is the probability of obtaining a result at least as extreme as the one that was actually observed (in our case 7 industries for which $m = 1$), given H_0. Let the collection of industries be I; that is, $I = \{1, 2, ..., 31\}$ is a set consisting of 31 elements. Let $M = \sum_{j \in I} m_j$. In our context,

$$p_7 = \Pr\left\{M \geq 7 \mid \left(\tau^j, N_X^j, N_E^j\right)_{j=1}^{31} \text{ and } H_0\right\}.$$

Of course, H_0 determines the ρ_j via (14). Let $\mathcal{J}_M$ be the collection of M-element subsets of I. That is, $\mathcal{J}_M = \{J \subset I \mid \#J = M\}$. Then,

$$p_n = \sum_{M \geq n} \sum_{J \in \mathcal{J}_M} \prod_{j \in J} \rho_j \prod_{j \notin J} \left(1 - \rho_j\right)$$

This is the likelihood of all possible values of M that equal n or more. The probability distribution of M under H_0 is $p_{n+1} - p_n$ and it is shown in Figure 9. Since $p_7 = 0.0217$, we reject H_0 at the 5% but not at the 1% level.

5 Capital replacement and exit

Our model assumes that $\varepsilon\left(\tau\right)$ jumps at the rate λ, but really only the first jump matters, since that first jump suffices to disqualify a firm from competing in the

market for new capital. The general question can be put as follows: 'Do incumbent firms develop a disadvantage in setting up machines and plants of the newest vintages and, if so, how quickly?' The literature offers some indirect evidence. Prusa and Schmitz (1994) find that a firm's initial product is better than its second product, its second better than its third, and so on. Christensen (1997) cites examples of disk-drive producers and printer producers that failed to invest in the new technologies that entrants brought in. Henderson (1993) finds that incumbents are less successful at research efforts to exploit radical innovation than entrants are. Supporting the vintage-capital model generally, Filson and Gretz (2004, Tables 1-5) support the proportion that 'new firms and industry laggards are the most likely pioneers of new product generations.' Their Tables 1-5 show that in the early days of a product (here a diameter of a disk), new entrants lead the pack in terms of quality (here defined as storage capacity). As the format matures, the incumbents become leaders. There are 5 new generations of drives in the data set; all of them were pioneered by spin-outs.[21]

We now show evidence that airline companies are more likely to exit when their capital is old. We use data on aircraft where accurate information on age is available. Among exiting firms, the age of the capital stock is always T.[22] Surviving capital is generally younger than that. If demand was almost perfectly inelastic, there would be very little investment between the spike dates, and the capital stock among surviving firms would be almost as old as that among exiting firms. On the other hand, if demand is highly elastic, most of the existing capital would be quite young. Thus if we let a^S be the average age of the surviving firms' capital stock, unless we specify the demand curve all we can say is that

$$0 < a^S < T. \tag{17}$$

Figure 10 shows the empirical counterparts of a^S (white dots) and T (black dots) for the U.S. airline industry, in which the airplanes of exiting airlines were on average 7 years older than those of the surviving airlines. After weighing by the number of observations, the difference is highly significant. Balloon size is for each series proportional to the square root of the number of observations, but the constant of

---

[21]The evidence is not unanimous, however. Dunne (1994) finds that the age of a plant is unrelated to the tendency to adopt advanced technology.

[22]The presence of an active market for leased planes in the airline industry leaves this implication largely intact. First, the competitive leasing price for a vintage-τ plane at date t would be $p(t)e^{g\tau} - c$ with airlines making zero profits at each instant. Our argument is that whether an airline leases or buys its equipment does not affect its ability to make use of it. An airline that has not experienced any jumps in ε can make use of the latest vintage at zero additional cost after paying the competitive rental. One that has experienced a shock to ε must also pay a positive amount in addition to the competitive rental to fly latest-vintage planes. But it can buy or rent planes of an older vintage, a vintage that it can still operate as efficiently as anyone else can. Suppose that the firm follows this strategy for as long as it can, thus hanging on for T periods following the shock to its ε. When it does finally exit, its rented planes will be T years old, i.e., the earliest vintage still in use in the industry, which also is when the equilibrium rental price of the vintage reaches zero.

Age of capital in years
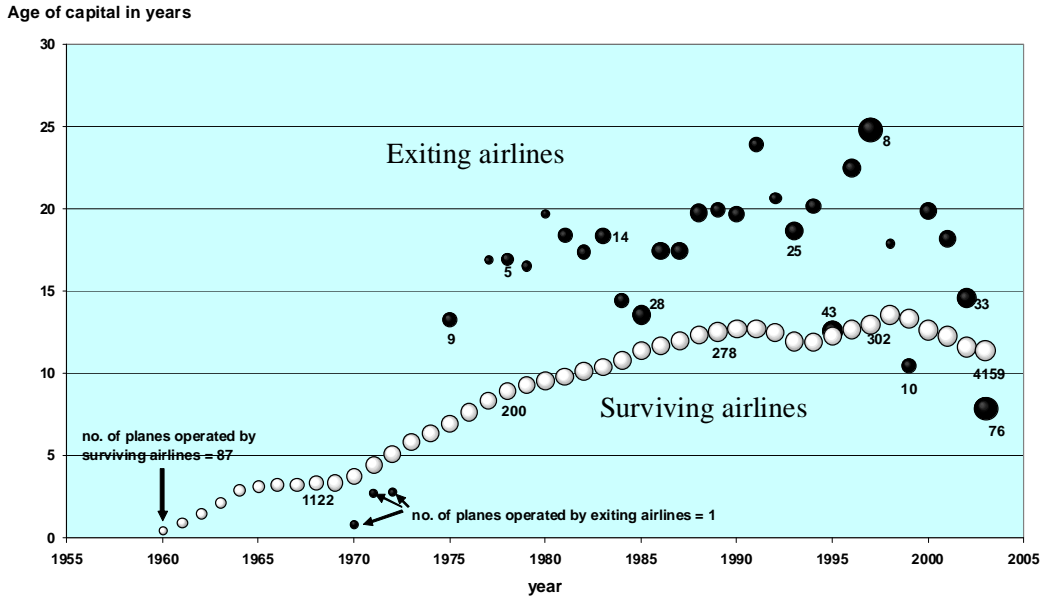
Figure 10: AGE OF CAPITAL AMONG EXITERS AND SURVIVORS IN THE U.S. AIRLINE INDUSTRY, 1960-2003.

proportionality is larger for the exiting capital series so as to allow us to see how sample size of that series too evolves over time.

Figure 11 shows almost as strong a difference among all the world airlines. Capital of exiting airlines was four years older than the capital of surviving airlines and the difference in the means is again highly significant. Surprisingly, perhaps, the U.S. carriers operated older planes than the carriers based in other countries. A description of the data is in the Appendix, but a summary is in Table 2:

	U.S.		**World**	
	Survivors	*Exiters*	*Survivors*	*Exiters*
Average age (yrs)	10.4	17	9.8	13.7
Std. dev. of age	7.5	9.4	7.4	9.4
# of airlines	231	181	1140	659
# of plane-year obs.	95,147	530	213,650	2,092

Table 2: THE DATA IN FIGURES 10 AND 11.

As reported in the table, a plane is counted as one observation for each year of its life. An airline is counted at most twice – as a survivor and then possibly as an exiter.

Three other pieces of evidence link exit decisions to the need to replace capital to the decision to exit.

Age of capital in years

Exiting airlines

12

35

2

24

10

406

871

300

97

2

552

7118

49

86

25

178

53

11169

Surviving airlines

year
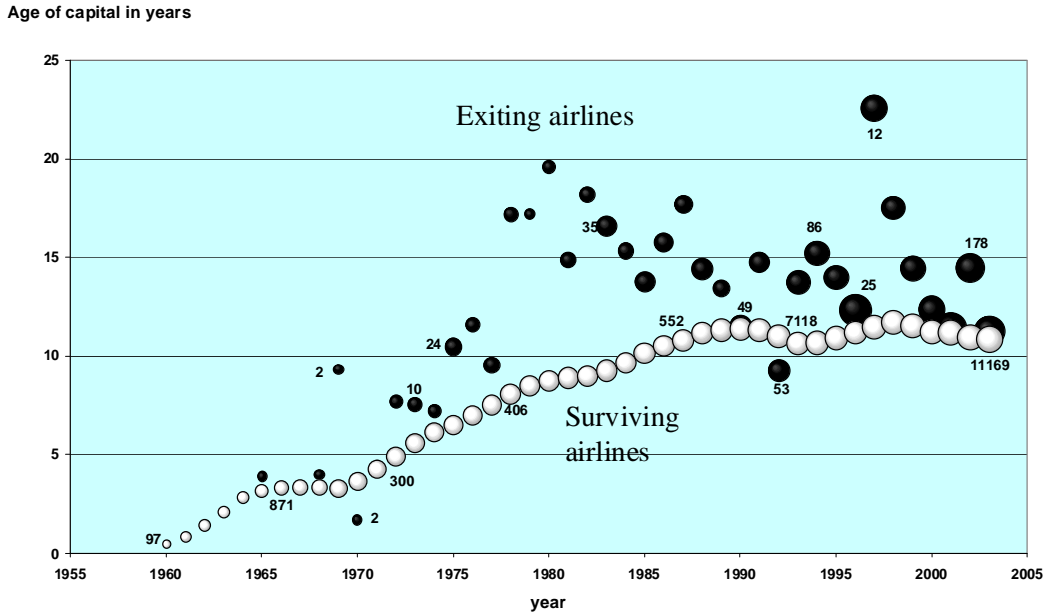
Figure 11: AGE OF CAPITAL OF EXITERS AND SURVIVORS IN AMONG ALL THE WORLD'S AIRLINES, 1960-2003.

1. *Plant exit.*—Salvanes and Tveteras (2004) find that old plants are (*i*) less likely to exit, but (*ii*) more likely to exit when their equipment is old. Fact (*i*) they attribute to the idea that plants gradually learn their productivity and exit if the news is unfavorable as Jovanovic (1982) argued. Fact (*ii*) they attribute to the vintage-capital effect on exit.

2. *The trading of patent rights.*—The renewal of a patent is similar to replacing a piece of capital in the sense that a renewal cost must be paid if the owner of the patent is to continue deriving value from it. If that owner sells the patent right, he effectively exits the activity that the patent relates to. Serrano (2008) finds that the probability of a patent being traded rises at its renewal dates, indicating that the decision to exit is related to the wearing out of a patent right.

3. *Higher embodied technical progress raises exit.*—A faster rate of decline in the price of capital makes it optimal to replace capital more frequently. If replacement sometimes leads to exit, we should see more exit where there is more embodied progress. Samaniego (*forthcoming*) indeed finds that in sectors where the prices of machinery inputs fall faster, the firms using those machines experience higher rates of exit.

23

5.0.3   Firm exit

A firm 'exits' in the GK data when it stops producing a product. In the theory presented so far, firms can exit only at ages $T, 2T$, and so on. Pooling over products, Agarwal and Gort (1996, Figure 2) show that hazard rates for *firm* exit tend to rise once age exceeds 30 or 40 years, and perhaps this is because their capital is old and needs replacing. It is conceivable that mixtures over industries could produce downward-sloping hazards. The fact is, however, that firms in every industry can and do fail at young ages and for reasons often unrelated to the age of their capital. Thus as a theory of exit within a particular industry, the model as it stands is quite inadequate. The object here is not to seek a general theory of firm exit but, rather, only to show that our theory of the shakeout is robust to the introduction of other reasons for exit.

We shall now show that the model can produce exit spikes every T periods, and still have a realistic exit hazard. To do it, we shall assume that firms experience cost shocks and that there is a frictionless used-capital market within the industry. This market becomes active if firms have different c's. We continue to assume that capital has zero value to anyone outside the industry. Suppose, then, that c rises permanently when the firm or plant is τ years old, and suppose further that the hazard, $h(\tau)$, of such an event is decreasing in τ, perhaps because of a positive effect that learning by doing exerts on the firm's production efficiency.

A firm that experiences a rise in c will immediately sell its capital to firms in the industry that have not experienced such a rise, and extract the full value of the profits that the machine will yield until its replacement date. The investment condition (5) is then unchanged because the resale value of the capital fully captures what the firm would have obtained for itself had its c remained at its original level. The new owner of the capital would choose to retire it at the same date that the original owner would have had he not experienced a rise in c and therefore the exit condition (4) also remains unchanged, as does (6). Equilibrium is therefore exactly the same. However, now there will be two kinds of exit: (i) When the firm's capital is up for replacement and that the firm's implementation cost has gone up since the last time capital was installed, and this hazard is zero except for spikes every T periods, (ii) When the firm's c has risen, which has the hazard $h(\tau)$. For $\tau < T$, the exit hazard is just h and it is decreasing in τ. At $\tau = T$, (i) dominates (ii) since time is continuous. But over a discrete time period, (ii) could dominate (i), in which case the exit hazard would decrease monotonically.[23]

The general point is that the GK observations that Figure 1 portrays concern industry aggregates. Our model too is mainly about these aggregates and it is probably consistent with a wide variety of assumptions about turbulence at the more micro level.

---

[23] This applies only to a firm that enters with frontier technology. If an entrant were to come in with used capital, he would replace it before his firm reached age T.

6 Other explanations

Our paper explains the existence of investment spikes using a vintage-capital argument. Proposition 3 goes further and predicts a geometric decay of the spikes which no other models generate in a natural way. On a general level, other forces have been held to be possible causes of concentrated episodes of investment and exit. Since our model omits these other forces, we now try to place our paper in a more general context by briefly reviewing these arguments and assessing their ability to explain the data at hand. It helps to divide them into demand-side and supply-side types of explanations.

6.1 Demand-side explanations

Some papers consider fluctuations in aggregate demand, others focus on shifts in demand at the industry level. We take these up in turn.

6.1.1 Aggregate demand

Gourio and Kashyap (2007) find investment spikes to be procyclical, and other work finds exits to be countercyclical. The concern this raises is that what we have termed as 'echoes' are instead reflections of the business cycle.

Be they aggregate or industry demand shocks, however, it would seem that they should produce an asynchronicity between entry and exit: Entry should occur when demand is high and exits should happen when demand is low. Our clustering test shows, however, that entry and exit spikes tend to go together. More directly we can ask if entry spikes are less likely and exit spikes more likely to occur during NBER recession years.[24]

This logic would seem to have been at work in the Phonograph industry in that, as Figure 6 shows, that in the two of the four exit spikes – the 1913 and the 1974 spikes – occurred during recession years and none of the entry spike dates did. However, this asymmetry does not show up in other industries. When we consider all the industries there is no correlation between recession years and spike dates. Our tests are simple comparisons of means. Let τ_j be the number of years of coverage of industry j, and τ_j^R the number of recession years contained in τ_j. Similarly, let ϕ_j be the number of

---

[24]The NBER dates are in quarters but our spike dates are in years. We designate a given year a recession year if any of the following criteria was met: (1) if a NBER recession was underway for the whole year, (2) if a NBER recession started in the first 2 quarters of the year, or (3) if a NBER recession ended in the last 2 quarters of the year. There are two exceptions. Both the 1918 and 1990 recessions started in the 3rd quater and ended in the 1st quater of the following year. They are nonetheless considered recession years, for otherwise the two recessions would not show in the data at all.

spikes in industry j and ϕ_j^R the number of those spikes that fall in a recession year.[25] Then we ask if either of the following inequalities is significant:

$$\frac{\sum_j \phi_j^R}{\sum_j \phi_j} \gtrless \frac{\sum_j \tau_j^R}{\sum_j \tau_j}.$$

The demand hypothesis implies that for exits the LHS should exceed the RHS and that for entry the LHS should be less than the RHS. It turns out that the LHS is smaller for both exit and entry, but neither difference is significant. We interpret this as a test of differences in the means based on $\sum_j \tau_j$ independent samplings. See Appendix Table 4.

6.1.2 Industry demand

We assumed that the demand curve is elastic enough and that large exogenous demand shifts are ruled out (the details are in (3)). When this assumption fails, demand can play a major role. This is shown by Caballero and Hammour (1994) who find that if replacement of capital causes interruptions in production, it is optimal to replace capital when demand temporarily drops, for then the foregone sales are at their lowest. Klenow (1998) finds that when, in addition, the productivity of new capital rises with cumulative output, a firm will replace capital when the recovery is about to start. This raises the productivity growth of the new capital.

If an exit spike is caused by a fall in industry demand, we should observe a fall in output around the time of the spike. Moreover, a demand decline produces a shakeout and a decline in p. If demand does not decline then there is no shakeout and no decline in p. We divide the GK industries into those in which output fell during the shakeout period, and into those in which output rose. For the analysis in Figure 4, among the industries where output fell during the shakeout,[26] the average pair was $\left(\widehat{g}, \widehat{T}\right) = (0.07, 30.4)$, whereas among the industries where output rose,[27] the average pair $\left(\widehat{g}, \widehat{T}\right) = (0.28, 13.3)$. For the analysis in Figure 5, the corresponding figures are $\left(\widehat{g}, \widehat{T}\right) = (0.05, 17.3)$[28] and $\left(\hat{g}, \widehat{T}\right) = (0.21, 11.2)$.[29] Yet the demand hypothesis

---

[25] Take for example product 23, Phonograph Records. Here $\tau = 83$. The coverage was from 1909 to 1991 during which 19 years were recession years. Thus $\frac{\tau_j^R}{\tau_j} = \frac{19}{83}$. For exit spikes, the second and the fourth spikes, in 1933 and 1973-74, respectively, were in recession years. Then $\phi_j^R = 2$ and $\frac{\phi_j^R}{\phi_j} = \frac{2}{19}$. Because of time aggregation, demand fluctuations can induce a spike in that $x(t)$ reacts positively to changes in demand.

[26] DDT, Electic shavers, Phonograph records, TV, and Zippers.

[27] Fluorescent lamps, Penicillin, and Styrene.

[28] The industries in note 26 plus Auto tyres and CRT.

[29] The industries in note 27 plus Freezers and Transistors. GK did not report whether output fell or rose during the shakeout for Computers and Streptomycin.

implies the opposite: Prompted by the decline in output, $\hat{g}$ (the measured *decline* in p) should have been higher and $\hat{T}$ should have been lower in the first sample than in the second. Thus demand shifts do not seem to explain the patterns in GK's sample.

6.2   Supply-side explanations

Echo effects were first discussed in the growth literature. In a one-consumption-good GE model, for the interest rate to remain constant in the face of variations in the rate of investment that inevitably occur along the transition path from arbitrary initial conditions, the instantaneous utility of consumption must be linear. Then, if the rate of technological progress is also constant, the investment echoes will have the constant periodicity that they also have in our model, but they will not be damped. Rather, the investment profile simply repeats itself every T periods – see Mitra *et al.* (1991) and BGL. Johansen (1959) and Arrow (1962) assume a production function for the sole final good that is Leontieff in capital and labor. In that case the effective maintenance cost is the wage multiplied by the labor requirement per machine. Since wages rise at the same rate as the rate at which the labor requirement declines, and the maintenance rate is then constant in units of the consumption good. Thus our maintenance cost, c, has an exact counterpart in these models. Since these models have no costs of rapid adjustment, the infinitely-elastic supply of capital that we assume amounts to the same thing.

Endogenous technological change.—Our model assumes that technological change is exogenous. Klepper (1996) assumes firms do research, and his model appears to imply that leading firms would squeeze out the inefficient fringe more quickly in industries where there is more technological opportunity and, hence, faster-declining product prices. The difference between our model and his concerns the fate of the firms in the first cohort of entrants: In our model the first cohort is the least efficient, whereas in Klepper's model, the first cohort is the *most* efficient because it has done the most research. In Grossman and Helpman (1991) and Aghion and Howitt (1992), monopoly incumbents are periodically replaced by more efficient entrants. Tse (2001) extends the model to allow for more than a single producer.

Exit after learning through experience.— GK observed that a shakeout usually comes after a wave of entry. Horvath, Schivardi and Woywode (2003, 'HSW') argue that if a run-up in entry occurs at some point in the industry's life, then some time later, a fraction of the entrants will have found themselves unfit to be in that industry – a type of learning that Jovanovic (1982) stresses – and will then exit *en masse*. The argument of HSW would measure the learning period by T. After exactly T periods, a firm discovers whether it has high costs and should exit. To develop this argument fully one would have to somehow to show first that a firm's learning about productivity is likely to take place all at once and not gradually as one usually assumes. And, second, one would need to explain why learning should occur sooner in those industries where p is declining rapidly. Nevertheless, HSW get some support

from the results in Table 7 of DRS where they find that at the industry level, the *number* of entering and exiting firms (as opposed to their market shares) is not positively correlated but that, rather, entry leads exit at a five-year frequency. In our data, however, the causality runs the other way: Exit appears to cause entry. For each industry, i, in Table 3 we define the dummy variables $x_{i,t}$ (exit) and $y_{i,t}$ (entry) taking on the value of unity if our procedure detects a spike at date t, and the value of zero if no spike is detected. Now we have a pair of sequences of dummy variables for each industry. We then run Probit regressions of the form

$$\Pr\left(x_{i,t} = 1\right) = \Phi\left(\alpha + \sum_{s=0}^{S} \alpha_s y_{i,t-s} + \sum_{s=1}^{S} \beta_s x_{i,t-s}\right) \tag{18}$$

and

$$\Pr\left(y_{i,t} = 1\right) = \Phi\left(\alpha + \sum_{s=0}^{S} \alpha_s x_{i,t-s} + \sum_{s=1}^{S} \beta_s y_{i,t-s}\right) \tag{19}$$

where Φ is the normal probability function. We pool the observations over (i, t). We report the results in Appendix, with S ranging from zero to five, and with and without restricting the coefficients. Generally, the contemporaneous relation is always positive – the estimate of α_0 is positive and significant at the 5% level in both (18) and (19) regardless of what other lags are included in them. The tendency is for exit to cause entry and not the other way around. This, however, was to be expected given the outcome of the coincidence test. Causality, on the other hand, tends clearly to run from exit to entry. Tables 5a and 5b tell the story in greater detail.

Technological advances by incumbents.—When a new technology raises the efficient scale of firms, it crowds some of them out of the industry; so argue Hopenhayn (1993) and Jovanovic and MacDonald (1994). The assembly-line technology, e.g., probably raised the optimal scale of auto-manufacturing plants and caused a large reduction in the number of auto producers. Klepper (1996) argues that larger firms do more R&D than small firms because they can spread its results over a larger number of units; because they invent at a faster rate, large firms then drive smaller firms out, often by acquiring them.

Consolidations for other reasons.—Some shakeouts no doubt occur because of the standardization of products so that some of the variants fall by the wayside and their producers exit. The focus on the diesel technology was one reason for the mass exodus of automobile producers that relied on other technologies as a source of power for the models they built. The winning model forces out other models and their producers. Utterback and Suarez (1993) argue that a dominant design emerges: More generally, consolidations and merger waves can occur for reasons unrelated to drops in demand and to advances in technology. Deregulation, for instance, has led to merger waves in the airlines and banking industries (Andrade *et al.* 2001) and to a sharp fall in the numbers of producers.

28

Evidence shows there are investment spikes – Cooper and Haltiwanger (1996), Gourio and Kashyap (2007). Repeated echo effects in entry and exit seem to be present in the data and they are clear indications that a vintage-capital replacement motive is one of their causes.

7   Conclusion

This paper started out with a graphical display of evidence that industry shake-outs of firms occur earlier in industries where technological progress is faster. We argued that other models of shakeouts were not able to explain this fact, whereas our vintage-capital model does so by predicting earlier replacement when capital-embodied technological progress is fast. We supported this claim with evidence from the airline industry that showed firm exits to be positively related to the age of the capital stock.

By inferring technological progress in the inputs from the decline in the price of the output as our model predicted, we found that the model fits fairly well the negative relation between technological progress and the onset of the shakeout. Moreover, we found that subsequent investment spikes, too, are also more frequent where technological progress is fast.

To study further the role of the vintage model in explaining industry dynamics, more can be done with richer data sets. First, the age of a firm's capital stock and the pressure that replacement of that stock places on the firm to exit should be linked empirically more firmly; we have studied only the airline industry. Second, the vintage-capital model that we have presented is primarily a theory of investment at the level of an industry, and could just as well be tested with data on industry-level gross and net investment; spikes should occur in the gross-investment series only, provided the quality of capital is properly measured. Third, we have abstracted entirely from mergers, yet horizontal mergers provide a way for exiting firms to be absorbed by those firms that remain; industry-level time series of horizontal mergers should, on these grounds, also exhibit echoes.

References

[1] Agarwal, Rajshree. "Evolutionary Trends of Industry Variables." *International Journal of Industrial Organization* 16, no. 4 (July 1998): 511-25.

[2] ———, and Michael Gort. "The Evolution of Markets and Entry, Exit and Survival of Firms." *Review of Economics and Statistics* 78, no. 3. (Aug., 1996): 489-98.

[3] Aghion, Philippe and Howitt, Peter, "A Model of Growth Through Creative Destruction," *Econometrica,* March 1992, 60(2), 323-351.

[4] Aizcorbe, Ana and Samuel Kortum. "Moore's Law and the Semiconductor Industry: A Vintage Model." *Scand. J. of Economics* 107, no. 4 (2005): 603–630.

[5] Andrade, Gregor; Mitchell, Mark and Stafford, Erik. "New Evidence and Perspective on Mergers." *Journal of Economic Perspectives,* Spring 2001, 15(2), pp. 103-120.

[6] Arrow, Kenneth. "The Economic Implications of Learning by Doing." *Review of Economic Studies* 29, no. 3 (June 1962): 155 - 173.

[7] Atkeson, Andrew and Patrick Kehoe. "Industry Evolution and Transition: The Role of Information Capital." FRB Minneapolis staff report 162 (August 1993).

[8] Boucekkine, Raouf, Marc Germain, and Omar Licandro. "Replacement Echoes in the Vintage Capital Growth Model." *Journal of Economic Theory* 74 (1997): 333-48.

[9] Caballero, Ricardo and Mohamad L. Hammour. "The Cleansing Effect of Recessions." *American Economic Review* 84, no. 5 (December 1994): 1350-68.

[10] Chatterjee, Satyajit and Rossi-Hansberg, Esteban. "Spinoffs and the Market for Ideas." FRB of Philadelphia Working Paper No. 08-26 (October 2008).

[11] Christensen, Clayton. *The Innovator's Dilemma*. Cambridge: HBS Press, 1997.

[12] Cooper, Russell and John Haltiwanger "Evidence on Macroeconomic Complementarities." *Review of Economics and Statistics* 78, no. 1 (February 1996): 78-93.

[13] Dunne, Timothy, Mark Roberts; Larry Samuelson. "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries." *RAND Journal of Economics* 19, no. 4 (Winter, 1988): 495-515.

[14] Dunne, Timothy. "Plant Age and Technology use in U.S. Manufacturing Industries." *RAND Journal of Economics* 25, no. 3 (Autumn, 1994): 488-499.

[15] Evans, David. "The Relationship Between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries." *Journal of Industrial Economics* 35, no. 4 (June 1987): 567-81.

[16] Filson, Darren and Richard Gretz. "Strategic Innovation and Technology Adoption in an Evolving Industry." *Journal of Monetary Economics* 51, no. 1 (January 2004): 89-121.

[17] Gavazza, Alessandro. "Leasing and Secondary Markets: Theory and Evidence from Commercial Aircraft", New York University, (February 2009).

[18] Gort, Michael and Steven Klepper "Time Paths in the Diffusion of Market Innovations," *Economic Journal* 92 (1982): 630-53.

[19] Gourio, Francois and Anil Kashyap. "Investment spikes: New facts and a general equilibrium exploration." *Journal of Monetary Economics* 54 (2007): 1–22.

[20] Grossman, Gene M. and Helpman, Elhanan, "Quality Ladders in the Theory of Growth," *Review of Economic Studies*, January 1991, 58(1), 43–61.

[21] Henderson, Rebecca. "Underinvestment and Incompetence as Responses to Radical Innovation: Evidence from the Photolithographic Alignment Equipment Industry." *RAND Journal of Economics* 24, no. 2 (Summer 1993): 248-270.

[22] Hopenhayn, Hugo. "The Shakeout." wp#33 Univ. Pompeu Fabra, April 1993.

[23] Horvath, Michael, Fabiano Schivardi and Michael Woywode. "On Industry Life-Cycles: Delay, Entry, and Shakeout in Beer Brewing." *International Journal of Industrial Organization* 19, no. 7 (2003): 1023-1052.

[24] Johansen, Leif. "Substitution versus Fixed Production Coefficients in the Theory of Economic Growth: A Synthesis." *Econometrica* 27, no. 2 (April 1959): 157-176.

[25] Jovanovic, Boyan. "Selection and the Evolution of Industry." *Econometrica* 50, no. 3 (June 1982): 649-70.

[26] ———, and Saul Lach "Entry, Exit, and Diffusion with Learning by Doing." *American Economic Review* 79, no. 4 (September 1989): 690-99.

[27] ———, and Glenn MacDonald. "The Life Cycle of a Competitive Industry." *Journal of Political Economy* 102 (April 1994): 322-47.

[28] ———, and Peter L. Rousseau. "General-Purpose Technologies." *Handbook of Economic Growth* vol. 1B Amsterdam: Elsevier (2006): 1181-224.

[29] Klenow, Peter. "Learning Curves and the Cyclical Behavior of Manufacturing Industries." *Review of Economic Dynamics* 1, no. 2 (May 1998): 531-50.

[30] Klepper, Steven, "Entry, Exit, Growth, and Innovation over the Product Life Cycle," *American Economic Review* 86 no. 3 (June 1996): 562 - 83.

[31] ———, and Kenneth Simons. "Industry Shakeouts and Technological Change." *International Journal of Industrial Organization* 23, no. 2-3 (2005): 23-43.

[32] ———, and Peter Thompson. "Disagreements and Intra-Industry Spinoffs." Florida International University working paper 09-07 (March 2009).

[33] McGrattan Ellen, and James A. Schmitz, Jr.. "Maintenance and Repair: Too Big to Ignore." *Federal Reserve Bank of Minneapolis Quarterly Review* 23, no. 4 (Fall 1999): 2–13.

[34] Mitchell, Matthew. "Technological Change and the Scale of Production." *Review of Economic Dynamics* 5, no. 2 (April 2002): 477-88.

[35] Mitra, Tappan, Debraj Ray and Rahul Roy. "The Economics of Orchards: An Exercise in Point-input, Flow-output Capital Theory." *Journal of Economic Theory*, 53, no. 1 (1993): 12-50.

[36] Prusa, Thomas and James Schmitz. "Are New Firms an Important Source of Innovation? : Evidence from the PC Software Industry." *Economics Letters* 35, no. 3 (March 1991): 339-342.

[37] Prusa, Thomas and James Schmitz. "Can Companies Maintain Their Initial Innovation Thrust? A Study of the PC Software Industry." *Review of Economics and Statistics* 76, no. 3 (August 1994): 523-40.

[38] Salvanes, Kjell and Ragnar Tveteras. "Plant Exit, Vintage Capital and the Business Cycle." *Journal of Industrial Economics* LII, no. 2 (June 2004): 255-76.

[39] Samaniego, Roberto M. "Entry, Exit and Investment-Specific Technical Change." *American Economic Review,* forthcoming.

[40] Serrano, Carlos. "The Dynamics of the Transfer and Renewal of Patents." NBER Working Paper 13938 (April 2008).

[41] Tse, Chung Yi, "The Distribution of Demand, Market Structure and Investment in Technology," *Journal of Economics*, October 2001, 73(3), 275–297.

[42] Utterback, James, and Fernando Suárez, "Innovation, Competition, and Industry Structure," *Research Policy* 22 (1993): 1-21.

[43] Wang, Zhu, "Learning, Diffusion and Industry Life Cycle." Federal Reserve Bank of Kansas City Working Paper 04-01 (January 2006).

#	Product & yr of comm intr.	Length of raw series	EXIT		ENTRY		Corr btw smoothed entry and exit rates	Corr btw entry and exit residuals	Corr btw entry and exit rates	G-K Stage 4
			spike dates		spike dates					
			in yrs since comm intro	in calendar yr	in yrs since comm intro	in calendar yr				
1	Antibiotics 1948	44	7	1955	2	1950	-0.46	-0.04	-0.30	not in G-K
			19	1967	39	1987				
			32	1980						
2	Artificial Christmas Trees 1938	54	8	1946	35	1973	-0.52	-0.02	-0.21	1968-1969
			18	1956	49	1987				
			43	1981						
			45	1983						
3	Ball-point Pens 1948	44	23	1971	9-10	1957-58	0.81	-0.32	0.55	S4 not reached
					26	1973-74				
4	Betaray Gauges 1956	36	7	1963	7*	1963	0.18	0.34	0.26	1973-
					33	1989				
5	Cathode Ray Tubes 1935	57	54	1989	10	1945	0.92	0.55	0.79	1963-1967
					15	1950				
					52	1987				
6	Combination Locks 1912	80	13	1925	21	1933	0.41	0.01	0.21	not in G-K
			16	1928	29	1941				
			65	1977	53	1965				
			75	1987						
7	Contact Lenses 1936	56	10	1946	6	1942	0.09	-0.15	-0.13	not in G-K
			29	1965	12	1948				
			35	1971	30	1966				
					39	1975				
8	Electric Blankets 1916	76	3	1919	6-7	1922-23	-0.13	0.07	-0.09	1962-1973
			35	1951	30	1946				
			41	1957	46	1962				
			63	1979	70*	1986				
			69-70	1985-86						
9	Electric Shavers 1937	55	36	1973	36-39*	1973-76	0.22	-0.09	-0.11	1938-1945
					49	1986				
10	Electrocardiographs 1942	50	6	1948	6*	1948	0.25	0.41	0.37	1964-1969
			32	1974	48	1990				
11	Freezers 1946	46	27	1973	40	1986	0.41	-0.13	0.23	1947-1957

italics - Is or may be in GK Stage 4   * - Within 1 year of exit spike

Table 3: Entry and Exit Statistics in Agarwal's data

	Product & yr of comm intr.	Length of raw series	EXIT		ENTRY		Corr btw smoothed entry and exit rates	Corr btw entry and exit residuals	Corr btw entry and exit rates	G-K stage 4
			spike dates		spike dates					
			in yrs since comm intro	in calendar yr	in yrs since comm intro	in calendar yr				
12	Freon Compressors 1935	57	5	1940	3	1938	-0.21	-0.12	-0.17	1971-1973
			37	*1972*	46	1981				
			52	1987						
13	Gas Turbines 1944	48	20	1964	2	1946	-0.42	0.00	-0.22	1973-
					36	1980				
					41	1985				
14	Guided Missiles 1951	41	2	1953	8	1959	-0.32	-0.43	-0.41	1965-1973
					35	1986				
15	Gyroscopes 1915	77	10	1925	4	1919	0.36	0.10	0.24	1966-1973
			31-32	1946-47	**30-32***	1945-47				
					39-40	1954-55				
16	Heat Pumps 1954	38	5	1959	25	1979	-0.42	-0.08	-0.24	1970-1973
			16	*1970*						
			32	1986						
17	Jet Engines 1948	44	6	1954	**6***	1953-54	-0.26	0.42	0.07	1960-1962
			22	1970	32	1980				
					35	1983				
18	Microfilm Readers 1940	52	5-6	1945-46	8	1948	0.37	-0.24	-0.12	S4 not reached
			16	1956	22-23	1962-63				
					33	1973				
					45	1985				
19	Nuclear Reactors 1955	37	30	1985	**30***	1985	-0.57	0.12	-0.38	1965-1973
			33	1988	35	1990				
20	Outboard Motors 1913	79	22	1935	2-4	1915-17	-0.18	0.07	-0.08	1921-1923
			67	1980	28	1941				
					34	1947				
					72	1985				
21	Oxygen Tents 1932	60	*41*	*1973*	4	1936	-0.32	-0.17	-0.26	1967-1973
			48	1980	56	1988				
			52-53	1984-85						
22	Paints 1934	58	*39*	*1973*	13-14	1947-48	-0.41	0.04	-0.20	1967-1973
			43	1977	20	1954				

italics - Is or may be in GK Stage 4   * - Within 1 year of exit spike

Table 3, continued

	Product & Yr of Comm Intr.	Length of raw series	EXIT		ENTRY		Corr btw smoothed entry and exit rates	Corr btw entry and exit residuals	Corr btw entry and exit rates	G-K stage 4
			spike dates		spike dates					
			in yrs since comm intro	in calendar yr	in yrs since comm intro	in calendar yr				
23	Phonograph Records	84	5	1913	9	1917	-0.23	0.08	-0.07	1923-1934
	1908		*25*	*1933*	13	1921				
			51	1959	40	1948				
			65-66	1973-74	**67***	1975				
24	Photocopying Machines	52	34	1974	42	1982	-0.30	-0.20	-0.25	1965-1973
	1940		39	1979						
25	Piezoelectric Crystals	51	13	1953	3	1943	-0.28	-0.06	-0.21	1955-1957
	1940		40	1980	45	1985				
26	Polariscopes	64	15	1943	18	1946	0.20	-0.26	-0.03	1964-1967
	1928		45	1973	23	1951				
					25	1953				
					44*	1972				
27	Radar Antenna Assemblies	40	*15*	*1967*	4	1956	-0.50	-0.10	-0.32	1957-1968
	1952		21	1973	33	1985				
28	Radiant Heating Baseboards	45	*25*	*1972*	27	1974	-0.42	0.02	-0.19	1972-1973
	1947		37	1984						
29	Radiation Meters	43			6	1955	-0.28	0.20	-0.13	not in G-K
	1949				37	1986				
			no spike dates		41	1990				
30	Recording Tapes	40	11	1963	33	1985	-0.53	0.06	-0.31	1973-
	1952		*28*	*1980*						
31	Rocket Engines	34	*32*	*1990*	7	1965	-0.34	-0.08	-0.19	1973-
	1958				**33***	1991				
32	Styrene	54	17	1955	4-5	1942-43	-0.55	-0.09	-0.27	1966-1973
	1938				20-1	1958-59				
33	Video Cassette Recorders	18			7	1981	-0.33	-0.18	-0.11	not in G-K
	1974		no spike dates		11	1985				
	average	51.94					-0.11	-0.01	-0.07	
	max	84.00					0.92	0.55	0.79	
	min	18.00					-0.57	-0.43	-0.41	
	std deviation	14.97					0.40	0.22	0.27	

italics - Is or may be in GK Stage 4   * - Within 1 year of exit spike

Table 3, continued

8   Appendix

8.1   The aircraft data

The data that Figures 10 and 11 portray come from a file, complied by a producer of a computer based aviation market information system and safety management software, that records the history of every non–military aircraft manufactured. They are described by Gavazza (2009, Sec. 5.1) and at *http://www.flightglobal.com/StaticPages/Acas.html* . Coverage begins in 1942 and ends at April 2003. For each aircraft, it records the usual identification information, such as the manufacturer, the model, etc., and (most important for our purposes) the first and the last date that each aircraft was operated by individual carriers. We aggregate such information at the level of an individual carrier to calculate the average age of capital of the carrier at a given point in time. The analysis is restricted to narrow-body (100 seats +) and wide–body passenger jets.

Measurement of entry and exit.—We say that a carrier has "entered" in the year in which it was first recorded as having operated a passenger jet. A carrier has "exited" in the year in which it was last recorded as having operated a passenger jet. A carrier has "survived" a certain year if it entered before that year and if it has not exited by the end of the year. The data contain information on mergers; a company that was acquired is not counted as an exit.

Measurement of age.—The age of a jet at a point in time is the time elapsed since the jet was first delivered. The average age of a carrier's capital is the average age of all the jets that it operated on January 1 of the year.

8.2   Proofs and derivations

Proof of Proposition 2.—Rewrite (9) as

$$ge^{-rT} + re^{gT} - \left(\frac{r+c}{c}\right)(r+g) \equiv \Phi\left(T, g, c, r\right) = 0.$$

We have $\frac{\partial \Phi}{\partial T} = rg\left(e^{gT} - e^{-rT}\right) > 0$. Therefore (11) follows if $\frac{\partial \Phi}{\partial g} > 0$. Now

$$
\begin{aligned}
\frac{\partial \Phi}{\partial g} &= e^{-rT} + rTe^{gT} - \frac{r+c}{c} \text{ and, by eliminating } e^{-rT}, \\
&= \left(\frac{r+c}{c}\right)\frac{r+g}{g} - \frac{r}{g}e^{gT} + rTe^{gT} - \frac{r+c}{c} \\
&= \left(\frac{r+c}{c}\right)\frac{r}{g} - \frac{r}{g}e^{gT} + rTe^{gT} = \frac{r}{g}\left(\frac{r+c}{c} - e^{gT}\right) + rTe^{gT} \\
&= \frac{r}{g}\left(\frac{r+c}{c} + (gT-1)e^{gT}\right) > 0.
\end{aligned}
$$

The strict inequality follows because (i) The function $(gT - 1)\,e^{gT}$ is increasing in gT, with derivative $e^{gT}\left[1 + (gT - 1)\right] = gTe^{gT}$, and because (ii) As $gT \to 0$, $(gT - 1)\,e^{gT} \to -1$, so that at its smallest point, $\frac{\partial \Phi}{\partial g} = \frac{r^2}{gd} > 0$. Then, since $\frac{\partial T}{\partial g} = -\frac{\partial \Phi}{\partial g}/\frac{\partial \Phi}{\partial T}$, and since $\frac{\partial \Phi}{\partial T} > 0$, (11) follows. ∎

Proof of Proposition 4.—(i) We showed that $\Phi_T > 0$ in the proof of Proposition 2. Moreover, $\frac{\partial \Phi}{\partial c} > 0$ because the ratio $\frac{r+c}{c}$ is decreasing in c. Therefore $\frac{\partial T}{\partial c} = -\frac{\partial \Phi}{\partial c}/\frac{\partial \Phi}{\partial T} < 0$. (ii) When $gc < r^2$, $\frac{\partial T}{\partial r} = -\frac{\partial \Phi}{\partial r}/\frac{\partial \Phi}{\partial T}$

$$
\begin{aligned}
\frac{\partial \Phi}{\partial r} &= -Tge^{-rT} + e^{gT} - \left(\frac{r+g}{c}\right) - \left(\frac{r+c}{c}\right) \quad \text{and, by eliminating } e^{gT}, \\
&= -Tge^{-rT} - \frac{ge^{-rT}}{r} + \left(\frac{r+c}{c}\right)\frac{(r+g)}{r} - \left(\frac{r+g}{c}\right) - \left(\frac{r+c}{c}\right) \\
&= -Tge^{-rT} - \frac{ge^{-rT}}{r} + \left(\frac{r+c}{c}\right)\frac{g}{r} - \left(\frac{r+g}{c}\right) \\
&= \frac{g}{r}\left(1 - e^{-rT}\right) - \left(\frac{r}{c} + Tge^{-rT}\right) < \frac{g}{r} - \frac{r}{c}.
\end{aligned}
$$

so that then $\frac{\partial T}{\partial r} = -\frac{\partial \Phi}{\partial r}/\frac{\partial \Phi}{\partial T} > 0$. ∎

Derivation of the Likelihood Function for the estimates in Figures 4, 5, 7, *and* 8—. For each industry i, if we know c_i and $\widehat{g}_i$, we know $\widetilde{T}_i = \widetilde{T}(r, c_i, \widehat{g}_i)$. We would like to solve $\widetilde{T}_i$ as a function of c_i, so that we can derive the distribution of $\widetilde{T}_i$ from the assumed log–normal distribution on c_i. But since $\widetilde{T}(r, c_i, \widehat{g}_i)$ does not admit a closed form solution, the exact functional relationship between $\widetilde{T}_i$ and c_i is unknown, and we linearize it in $x = \ln c$. After substituting $x_i \equiv \log c_i$ and $g = \widehat{g}_i$ into (9), it reads

$$
\left(\frac{r}{e^{x_i}} + 1\right)(r + \widehat{g}_i) = \widehat{g}_i e^{-r\widetilde{T}_i} + re^{\widehat{g}_i \widetilde{T}_i}. \tag{20}
$$

Taking total derivatives at $x_i = \overline{x}$, gives

$$
-re^{-\overline{x}}(r + \widehat{g}_i)\,dx_i = r\widehat{g}_i\left(e^{\widehat{g}_i \overline{T}_i} - e^{-r\overline{T}_i}\right)dT_i
$$

where $\overline{T}_i \equiv \widetilde{T}\left(r, e^{\overline{x}}, \widehat{g}_i\right)$. Then

$$
-\frac{dT_i}{dx_i} = e^{-\overline{x}}\frac{r + \widehat{g}_i}{\widehat{g}_i}\frac{1}{e^{\widehat{g}_i \overline{T}_i} - e^{-r\overline{T}_i}} \equiv \beta_i
$$

Hence $T_i \simeq \widetilde{T}(r, \overline{x}, \widehat{g}_i) - (x_i - \overline{x})\beta_i$. If $x_i \sim N(\overline{x}, \sigma^2)$, then approximately, $T_i \sim N\left(\overline{T}_i, \beta_i^2 \sigma^2\right)$. The density of T_i is $f(T_i) = \frac{1}{\sqrt{2\pi}\beta_i \sigma} \exp\left(-\frac{1}{2}\left[\frac{T_i - \overline{T}_i}{\beta_i \sigma}\right]^2\right)$. When there are no censored observations, the likelihood is

$$
L = \prod_{i=i}^{n} f_{T_i}(T_i);
$$

37

otherwise
$$L = \prod_{i \notin \mathcal{C}} f_{T_i}(T_i) \prod_{i \in \mathcal{C}} (1 - F_{T_i}(T_i)),$$

where $\mathcal{C}$ is the set of censored observations and $F_{T_i}(T_i)$ the CDF of $f(T_i)$.

Conditional Expectation in Figure 5.—Take the Nylon observation for example. The observation is truncated at 21. As we showed in the previous paragraph, $T_i \sim N\left(\bar{T}_i, \beta_i^2 \sigma^2\right)$ approximately. By setting $\bar{x}$ and σ equal to their respective point estimates of -2.56 and 2.27, we obtain estimates of $\bar{T}_i$ and $\beta_i^2 \sigma^2$. Then we calculate the expected value of T_i conditional on $T_i \geq 21$.

8.3 Tests

8.3.1 NBER recessions tests

The dates come from http://www.nber.org/cycles/cyclesmain.html. Table 4 summarizes our findings on the incidences of exit and entry spikes during the recessions:

		Exit Spikes		Entry Spikes	
$\sum_j \tau_j$	1681	$\sum_j \phi_j$	72	$\sum_j \phi_j$	95
$\sum_j \tau_j^R$	364	$\sum_j \phi_j^R$	14	$\sum_j \phi_j^R$	15
$\frac{\sum_j \tau_j^R}{\sum_j \tau_j}$	0.22	$\frac{\sum_j \phi_j^R}{\sum_j \phi_j}$	0.19	$\frac{\sum_j \phi_j^R}{\sum_j \phi_j}$	0.16

Table 4: SPIKES AND RECESSIONS

For both exit and entry spikes then $\sum_j \phi_j^R / \sum_j \phi_j < \sum_j \tau_j^R / \sum_j \tau_j$. The differences however are not significant at even the 10% level in chi-square tests of independence in both cases.

8.3.2 Causality tests

The tables below present the results of Granger causality tests of exit by entry (Table 5a) and of entry by exit (Table 5b). All regressions are Probit regressions. In each set of results, the first table presents the results where the coefficients are not restricted, whereas the second table presents the results where the coefficients of the various lags of each variable are restricted to be equal. Note how the results in general support exit causing entry but not the other way around.

Dep var = x_{it}

S	0	1	2	3	4	5
Constant	**-1.75***	**-1.74***	**-1.71***	**-1.69***	**-1.65***	**-1.66***
	(0.057)	(0.059)	(0.061)	(0.064)	(0.065)	(0.069)
y_{it}	**0.37**[†]	**0.38**[†]	**0.40**[†]	**0.41**[†]	**0.44**[†]	**0.46**[†]
	(0.193)	(0.196)	(0.199)	(0.201)	(0.204)	(0.206)
y_{it-1}		-0.25	-0.24	-0.23	-0.24	-0.17
		(0.268)	(0.277)	(0.278)	(0.282)	(0.289)
y_{it-2}			-0.101	-0.07	-0.07	-0.24
			(0.247)	(0.238)	(0.259)	(0.327)
y_{it-3}				**-0.56**[‡]	**-0.53**[‡]	**-0.49**[‡]
				(0.383)	(0.386)	(0.390)
y_{it-4}					**-0.57**[‡]	**-0.55**[‡]
					(0.391)	(0.394)
y_{it-5}						-0.08
						(0.265)
x_{it-1}		0.23	0.23	0.22	0.20	0.22
		(0.236)	(0.237)	(0.238)	(0.238)	(0.238)
x_{it-2}			**-0.51**[‡]	**-0.52**[‡]	**-0.55**[‡]	**-0.52**[‡]
			(0.394)	(0.395)	(0.398)	(0.400)
x_{it-3}				-0.15	-0.18	-0.16
				(0.311)	(0.313)	(0.325)
x_{it-4}					-0.20	-0.19
					(0.319)	(0.325)
x_{it-5}						-0.19
						(0.322)

Dep var = x_{it}

S	0	1	2	3	4	5
Constant	**-1.75***	**-1.96***	**-1.68***	**-1.61***	**-1.55***	**-1.57***
	(0.057)	(0.226)	(0.108)	(0.233)	(0.254)	(0.148)
$\sum_{s=0}^{S} y_{it-s}$	**0.37**[†]	0.12	0.06	-0.20	-0.08	-0.08
	(0.193)	(0.139)	(0.118)	(0.120)	(0.123)	(0.101)
$\sum_{s=1}^{S} x_{it-s}$		0.22	-0.03	-0.07	-0.10	-0.10
		(0.208)	(0.087)	(0.202)	(0.216)	(0.116)

* significant at the 1% level
[†] significant at the 5% level
[‡] significant at the 10% level

Table 5a: Causality Tests: Entry causes exit

Dep var = y_{it}

S	0	1	2	3	4	5
Constant	-1.61*	-1.67*	-1.68*	-1.70*	-1.73*	-1.74*
	(0.051)	(0.056)	(0.059)	(0.062)	(0.067)	(0.069)
x_{it}	0.39†	0.37†	0.40†	0.41†	0.44†	0.48†
	(0.203)	(0.207)	(0.208)	(0.210)	(0.210)	(0.211)
x_{it-1}		0.21	0.20	0.24	0.27	0.28
		(0.22)	(0.220)	(0.222)	(0.223)	(0.222)
x_{it-2}			0.22	0.20	0.25	0.25
			(0.220)	(0.227)	(0.225)	(0.228)
x_{it-3}				0.45†	0.46†	0.48†
				(0.207)	(0.208)	(0.209)
x_{it-4}					0.32‡	0.32‡
					(0.221)	(0.222)
x_{it-5}						0.32‡
						(0.225)
y_{it-1}		0.68*	0.71*	0.68*	0.65*	0.61*
		(0.164)	(0.167)	(0.171)	(0.177)	(0.183)
y_{it-2}			-0.21	-1.53	-0.23	-0.23
			(0.245)	(0.244)	(0.270)	(0.269)
y_{it-3}				-0.47‡	-0.43‡	-0.42‡
				(0.322)	(0.319)	(0.322)
y_{it-4}					-0.11	-0.13
					(0.266)	(0.255)
y_{it-5}						-0.02
						(0.148)

Dep var = y_{it}

S	0	1	2	3	4	5
Constant	-1.61*	-2.34*	-2.01*	-1.84*	-1.78*	-1.76*
	(0.051)	(0.190)	(0.156)	(0.131)	(0.137)	(0.084)
$\sum_{s=0}^{S} x_{it-s}$	0.39†	0.29†	0.28†	0.33*	0.37*	0.38*
	(0.203)	(0.146)	(0.122)	(0.106)	(0.097)	(0.091)
$\sum_{s=1}^{S} y_{it-s}$		0.67*	0.33*	0.14‡	0.05	0.02
		(0.165)	(0.126)	(0.100)	(0.099)	(0.042)

* significant at the 1% level
† significant at the 5% level
‡ significant at the 10% level

Table 5b: Causality Tests: Exit causes entry