NBER WORKING PAPER SERIES

THE ECONOMICS OF HAS-BEENS

Glenn MacDonald
Michael Weisbach

Working Paper 8464
http://www.nber.org/papers/w8464

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2001

This paper was prepared for the Conference in Honor of Sherwin Rosen. For thoughtful input, we thank Rui Albuquerque, Gary Becker, Ed Prescott, and especially Bruce Weinberg. An earlier version circulated under the title "Technology-specific capital and the success and failure of firms." The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

The Economics of Has-Beens
Glenn MacDonald and Michael Weisbach
NBER Working Paper No. 8464
September 2001
JEL No. D2, L2, G3

ABSTRACT

Evolution of technology causes human capital to become obsolete. We study this phenomenon in an overlapping generations setting, assuming it is hard to predict how technology will evolve, and that older workers find updating uneconomic.

Among our results is the proposition that (under certain conditions) a more rapid pace of technological advance is especially unfavorable to the old in the sense that the implied within-industry division of output or income between young and old becomes much more skewed, i.e., a smaller number of young earn comparatively more. We apply our results to architecture, an occupation in which the has-beens phenomenon has had a particularly acute impact.

Glenn MacDonald
Olin School of Business
Washington University in St. Louis

Michael Weisbach
College of Commerce and Business Administration
University of Illinois
and NBER

1 Introduction

Human capital investments are risky. Complementarities among skills, both in how they are learned and how they are used, imply that an efficient human capital portfolio consists of skills whose values are highly positively correlated. One of the main risks with which owners of human capital must contend is technological obsolescence. Best practice methodology is always evolving, and may do so in ways that preserve the value of some skills while depreciating others dramatically. Architecture is a good example of the latter.[1] Advances in computing have revolutionized the field, increasing the art component in what had traditionally been a blend of art, engineering and drafting. This lowered the cost of design development drastically, and altered the cost of experimentation or modification of plans even more. Older architects have found it uneconomic to master the complex computer skills that enable the young to produce architectural services so easily and flexibly, and have found hiring a "computer department" to be a poor substitute. Thus, these advances have allowed younger architects to serve much of the market for architectural services, causing the older generation to lose much of its business. For example, according to the 1990 Census of Population and Housing, white architects (engineers and surveyors are included here too) with Bachelor's degrees, aged 25-34, earned an average of more than $130,000 in 1989, whereas those aged 45-54 earned just over $100,000. The corresponding numbers for lawyers and judges, for whom practice and experience fosters understanding of new laws, court decisions, etc. are roughly

[1]We thank Thomas Montalto, of Montalto Architecture, for providing information on the history and impact of digital technology in architecture.

$42,000 and just over $82,000.

In this paper we develop a simple model of technology-specific human capital investment, and study the impact of unforeseeable technology evolution. The model has an infinite-horizon, overlapping generations, structure in which agents, when young, make human capital investments that are specific to the current cutting edge technology; old workers do not reinvest.[2] As technology evolves, the value of the human capital of older workers is eroded – to varying degrees depending on the extent of technology improvement – by competition from young workers whose skills are tailored to the newest technology. We characterize equilibrium entry of young workers, especially the degree to which advances in technology generate entry. Under certain conditions, entry displays a feature similar to one that is well-known to students of industry life cycles. There, new technology leads to a smaller number of large scale producers, with the less efficient being driven out; see Gort and Klepper (1982). In our model, the less efficient are older workers whose sunk investments lead them to continue to work, and technology advance results in a comparatively small number of young, highly productive, workers; reduced entry substitutes for exit by the inefficient. And this is more so the greater is the growth in technology. Thus, when technology improves more, the within industry distribution of income between young and old becomes much more skewed in that a smaller number of younger workers earn comparatively more. Older workers, who are invariably has-beens to some extent,

[2]Zeckhauser (1968) and Parente (1994) have studied optimal skill updating when there is learning-by-doing given use of a particular technology, but technology is steadily improving. See also Greenwood and Jovanovic (forthcoming).

are especially impacted when technology evolves rapidly since they are comparatively more numerous, earn comparatively less, and are absolutely worse off in that the price of their output is lower.

The has-been phenomenon is different from Rosen's (1981) superstars idea, but has clear similarities. Superstars earn relatively more as a result of superior talent and the way talent interacts with other choices. There is free entry to being ordinary, and superior ability commands a rent. In the dynamic, rising stars, version, the young enter hoping that as time passes events will reveal superior talent, and that superstardom will follow; see MacDonald (1988). In the has-beens phenomenon, there is free entry to use of advanced technology, and within-industry differences in relative output or compensation reflect deterioration in rents earned by those whose human capital has been depreciated in the economic, as opposed to physical, sense; Cooley et. al. (1997) and Campbell (1998) emphasize the same distinction with respect to physical capital.[3] The young enter hoping that technology will evolve in a way that depreciates their skills minimally as they age.

The interaction of technology-specific capital and technology evolution is a much more general phenomenon than the one we study here. Similar ideas apply to investment in physical capital, or acquisition of human capital inside organizations. For example, historically, many companies have encouraged workers to develop firm-specific human capital by structuring workers' compensation so that they have incentives to remain with the orga-

[3]Strictly, entry to use of new technology is not exactly free, since we assume older workers do not reinvest. It is easy to extend the model to allow for reinvestment, in which case this discussion applies to equilibria where the old do not reinvest.

nization. Thus, steeply-sloped age-earnings profiles, defined-benefit pension plans containing implicit penalties for changing jobs, and even promises of "lifetime" employment have been common at firms where human capital is particularly important; see, e.g., Lazear (1995). A technological shift that obsoletes the human capital of most of the workers leaves the firm with the unenviable choice of proceeding with the same work force against competitors trained in the better technology, or laying off workers, breaking implicit promises to them and living with the ensuing reputational damage. Eastman Kodak found itself in such a position in the early 1990s when the state of the art in the image-processing industry shifted from a chemical approach to an electronic one. And Xerox is facing the same challenges today as new approaches to reproduction, storage and dissemination of documents were based on digital, instead of analogue, technologies.

In the Sections following, we set out the model, define an equilibrium, and discuss our results. Then we discuss some extensions, limitations, and related ideas, e.g., the model suggests a novel explanation for academic tenure. All proofs appear at the end.

2 Model

Time is discrete, and the horizon infinite; $t = 0, 1, ...$ At each t a unit mass of agents is born and lives for two periods. Thus there is a continuum of "young" (y) and "old" (o) agents at every date beyond $t = 0$. All agents are risk neutral expected wealth maximizers and inelastically supply a fixed amount of time to work each period. Denote their common discount factor by

$\beta; \beta \in (0,1)$. To focus on the impact of accumulation of technology-specific capital, we assume agents are born identical.

The economy comprises an industry (defined by the homogeneous, non-storable product firms produce) that is small in relation to the rest of the economy. The impact of the small industry assumption is that the evolution of the rest of the economy is independent of the way technology evolves in the industry.

Demand for the industry's product is stationary and depends on the current price of the product (p) and aggregate income (Y); the numeraire is the composite good produced elsewhere in the economy. Inverse demand is $p = D(Y, Q)$,where Q is quantity, and D is a positive, bounded, continuous and increasing function of Y/Q, specifically,

$$p = D(Y/Q). \tag{1}$$

Agents decide whether to work in the industry or in the rest of the economy. Whichever is chosen, the agent, when young, must make a given technology-specific human capital investment. This capital may be put to work immediately (i.e., when the agent is young) and old agents cannot reinvest. Also, human capital is useless unless employed in conjunction with the associated technology, and there is no other human capital. Direct investment costs are ignored – investment costs are all opportunity costs. The impact of these assumptions is that old agents will do what they did when they were young. More generally, allowance for old agents' "retooling" or retiring early can be accommodated.

The technology-specific investment can be given a broad interpretation. For example, while it is convenient to refer to it as human capital, the model can equally be interpreted as physical or even organizational capital in the sense of a way of doing things, corporate culture, etc. The key is that having made the investment, it will be in place and in use for a period of time exceeding that for which the current technology is expected to be at the cutting edge.

Having made a technology-specific investment, the agent can put it to use. Production is modeled as an individual activity, although, as usual, it can be interpreted as the outcome of constant returns to scale production. For now, other factors of production are ignored; but see Subsection 4.5.2.

Let productivity (i.e., what one agent can produce in one period) in the rest of the economy be q_a. This technology improves by the fixed factor $\delta \geq 1$ every period; assume $\beta\delta < 1$. Thus, denoting future values of variables with a "$'$", $q_a' = \delta q_a$. Note that a young agent operating in the rest of the economy earns q_a for two periods as opposed to q_a when young and δq_a when old. The interpretation of this specification is that productivity growth is the result of technological advance as opposed to the accumulation of other capital inputs. The opposite assumption can be made with trivial changes to the analysis.

A young agent who elects to produce in the industry invests in the best practice technology and, as explained above, puts it to use for two periods. Let q_y denote the productivity of that technology. Productivity q_y is accessible to all entrants. The productivity of old agents (q_o) is last period's best practice productivity; i.e., $q_o' = q_y$.

The idea that the evolution of technology in the industry is both hard to predict and to some degree distinct from the evolution of technology in the rest of the economy plays a central role in the model. At this point (more structure will be added below) it suffices to assume that

$$q_y' \sim Q(\cdot \mid q_y, q_a) \tag{2}$$

where Q is a distribution function satisfying $Q(q_y' \mid q_y, q_a) = 0$ if $q_y' < q_y$.

3 Analysis

3.1 Agent optimization

Young agents must decide whether to enter the industry or to participate in the rest of the economy. This decision turns on a comparison of expected discounted income. To make this comparison, let $\tilde{s} \equiv (q_a, q_y, q_o, n)$ be an aggregate state vector summarizing opportunities in the industry and elsewhere, q_a and q_y, as well as the productivity and mass of old workers in the industry, q_o and n. Given $\tilde{s}$, let $P(\tilde{s})$ be the prevailing product price, and $e(\tilde{s})$ the mass of young entrants. Agents take P and e as given.

Consider a young worker's decision problem given some $\tilde{s}$. Since old workers use the pre-existing best technology, and so earn income pq_o when the price is p, maximum lifetime expected income for a young worker is given by

$$\max \left\{ q_y \left[P(\tilde{s}) + \beta \int P(\tilde{s}') dQ(q_y' \mid q_y, q_a) \right], q_a(1+\beta) \right\},$$

where $q_a' = \delta q_a, q_o' = q_y$, $q_y' \sim Q(\cdot \mid q_y, q_a)$ and $n' = e(q_a, q_y, q_o, n)$. The first component in the max{} is the discounted value of lifetime income the agent expects to earn if participating in the industry, comprising earnings while young, $P(\tilde{s})q_y$, and earnings expected while old, $q_y \int P(\tilde{s}')dQ(q_y' \mid q_y, q_a)$; the pace of technological progress is the source of uncertainty about the future.

The other component is the value the agent will certainly earn by investing in the rest of the economy and producing q_a in each period.

3.2 Equilibrium

Since all young workers are identical, equilibrium restricts the mass of young agents entering the industry, but not the identities of the entering workers. This is analogous to the familiar perfectly competitive model of homogenous firms in which entry must suffice to ensure zero profit for entering firms, but entry by any set of firms yielding this outcome is consistent with equilibrium. Thus, equilibrium is defined in terms of the mass of entrants in any state. Further, the equilibrium price must equate supply and demand for the industry's product.

Definition 1 An *equilibrium* is a pair of functions, $(e(\cdot), P(\cdot))$ such that for all $\tilde{s}$

$$q_y \left[P(\tilde{s}) + \beta \int P(\tilde{s}')Q(dq_y' \mid q_y, q_a) \right] \begin{cases} = q_a(1+\beta) \\ \quad \text{if } e(\tilde{s}) > 0 \\ \\ \leq q_a(1+\beta) \\ \quad \text{if } e(\tilde{s}) = 0, \end{cases} \tag{3}$$

and

$$P(\tilde{s}) = D\left(\frac{q_a(1+\frac{1}{\delta})}{nq_o + e(\tilde{s})q_y} \right). \tag{4}$$

According to (3), if a positive mass of agents enter the industry, equilibrium requires that their expected discounted lifetime income equal the value they could achieve elsewhere in the economy. Consistent with the assumption that the industry is small, we suppose young agents are abundant enough to allow this condition to be met for any $\tilde{s}$. Likewise, if there is no entry, expected discounted lifetime income must be no greater than the value an agent could achieve elsewhere. Equation (4) requires that the product market clear in each state. Since the industry is assumed small relative to the economy as a whole, aggregate income is simply total income earned elsewhere in the economy, $Y = q_a(1+\frac{1}{\delta})$. Also, industry output is $Q = nq_o + e(\tilde{s})q_y$. Equation (4) is simply (1) with these substitutions.

3.3 Some Simplification

To proceed, it is convenient to observe that equilibrium entry maximizes the present discounted value of consumers' plus producers' surplus. Thus, properties of equilibrium entry can be derived by studying the policy function of the dynamic program corresponding to surplus maximization.

Lemma 1 $(e(\cdot), P(\cdot))$ is an equilibrium if and only if (4) is satisfied and $e(\cdot)$ is the policy function associated with the dynamic program

$$\widetilde{W}(\tilde{s}) = \max_{e \geq 0} \left\{ \int_0^{nq_o+eq_y} D\left(\frac{q_a(1+\frac{1}{\delta})}{x}\right) dx - eq_a(1+\beta) + \beta \int \widetilde{W}(\tilde{s}')Q(dq_y' \mid q_y, q_a) \right\}, \tag{5}$$

where $q_a' = \delta q_a, q_o' = q_y$, $q_y' \sim Q(\cdot \mid q_y, q_a)$ and $n' = e(q_a, q_y, q_o, n)$.

The intuition for this result is straightforward. Consider the surplus consequences of increasing entry slightly, say by $\Delta e > 0$. Since each new entrant produces q_y in each period, output rises by $q_y\Delta e$ in each period. The extra surplus created in the industry, in units of the numeraire, is then $q_y\Delta e P(\tilde{s})$ in the current period, and (in expectation) $q_y\Delta e \int P(\tilde{s}')Q(dq_y' \mid q_y, q_a)$ later.

The cost of the incremental entry is the current and future value of the forgone output elsewhere, $q_a(1+\beta)\Delta e$. Thus, if surplus maximization dictates strictly positive entry, the optimal entry level equates marginal returns and costs:

$$q_y P(\tilde{s}) + \beta q_y \int P(\tilde{s}')Q(dq_y' \mid q_y, q_a) = q_a(1+\beta);$$

otherwise

$$q_y P(\tilde{s}) + \beta q_y \int P(\tilde{s}')Q(dq_y' \mid q_y, q_a) \leq q_a(1+\beta),$$

i.e., together, (3).[4]

Next, there are two features of the environment that are especially important for the young agent contemplating entry. First, how much output is produced by the old workers in the industry? In equilibrium, the old workers' output will influence the current equilibrium price and, thus, how valuable entering the industry will be. Note that it is not the mass of old workers, n, that matters; instead, it is their aggregate output, nq_o. Second, how productive will an entrant be, i.e., q_y? These variables, expressed in in comparison to a potential entrant's productivity elsewhere – i.e., $\frac{nq_o}{q_a}$ and $\frac{q_y}{q_a}$ – are important factors determining the value of entry relative to the value of working in the rest of the economy. Since the analysis of equilibrium entry is both facilitated and made more intuitive by focusing on $\frac{nq_o}{q_a}$ and $\frac{q_y}{q_a}$, we add modest additional structure sufficient to do so. Specifically, using (4),

[4] It would be more conventional to write the Bellman equation as

$$\widetilde{W}(\tilde{s}) = \max_{e\geq 0}\left\{\int_0^{nq_o+eq_n} D\left(\frac{q_a(1+\frac{1}{\delta})}{x}\right)dx - n\frac{q_a}{\delta} - eq_a + \beta\int \widetilde{W}(\tilde{s}')Q(dq_n' \mid q_n, q_a)\right\}.$$

Since the evolution of q_a is deterministic, this is equivalent to the more convenient formulation provided in the text.

making the change of variable $y = \frac{x}{q_a}$, and rearranging a bit, (5) becomes

$$\frac{\widetilde{W}(\tilde{s})}{q_a} = \max_{e \geq 0} \left\{ \int_0^{e\frac{q_y}{q_a} + n\frac{q_o}{q_a}} D\left(\frac{1+\delta}{\delta y}\right) dy - e(1+\beta) + \beta\delta \int \frac{\widetilde{W}(\tilde{s}')}{q_a'} Q(dq_y' \mid q_y, q_a) \right\}. \tag{6}$$

Now define

$$\xi \equiv \frac{q_y}{q_a},$$

i.e., productivity in the industry in units of productivity elsewhere, and assume that Q is such that

$$\xi' \sim \Xi(\cdot \mid \xi), \tag{7}$$

where $\Xi(\cdot \mid \xi)$ is (weakly) stochastically increasing in ξ; i.e., growth in technological know-how is persistent to some degree. As an example of a conditional distribution of q_y' consistent with (7), suppose

$$q_y' = X \max\{q_y, q_a\},$$

where X is random variable with support $[1, \bar{x}], \bar{x} < \infty$. Then, dividing by q_a and recalling $q_a' = \delta q_a$,

$$\xi' = \frac{X}{\delta} \max\{\xi, 1\},$$

and the associated distribution function satisfies (7).[5]

Next, define

$$\zeta \equiv \frac{nq_o}{q_a}$$

to be the total productivity of old workers in the industry in units of productivity elsewhere; ζ measures the "installed base" of the existing technology

[5]We assume Ξ has a positive and compact support containing unity in its interior, is stochastically increasing in ξ, and has the Feller property (i.e., if f is a bounded and continuous function of ξ, then so is $\int f(\xi')\Xi(d\xi' \mid \xi)$).

at any point in time. Observe that

$$\zeta' = \frac{n' q_o'}{q_a'} = \frac{e q_y}{\delta q_a} = \frac{e\xi}{\delta}.$$

Then, letting

$$z \equiv \frac{e\xi}{\delta}$$

be *next period's* installed base, and recalling that $q_a'/q_a = \delta$, we can rewrite the dynamic programming problem in terms of the two state variables ξ and ζ:[6]

$$W(\xi, \zeta) = \max_{z \geq 0} \left\{ \int_0^{\zeta + \delta z} D\left(\frac{1+\delta}{\delta y}\right) dy - \frac{\delta z}{\xi}(1+\beta) + \beta\delta \int W(\xi', z)\Xi(d\xi' \mid \xi) \right\}. \tag{8}$$

Let $Z(\xi, \zeta)$ be the optimal z given any (ξ, ζ), i.e., the policy function associated with (8). Note that given ξ and δ, varying z involves varying *current* entry, e.

It is easy to show that the value function W is increasing in both arguments. We also assume that W is strictly concave and continuously differentiable in ζ. In the context of our model, assumptions guaranteeing this are

[6]Let $S(s) \equiv \{\tilde{s} \mid \frac{q_n}{q_a} = \xi, \frac{n q_o}{q_a} = \zeta\}$, and observe that if $\widetilde{W}_0(\tilde{s})$ is any function that, for each s, is constant on $S(s)$, the function $\widetilde{W}_1(\tilde{s})$ defined by

$$\widetilde{W}_1(\tilde{s}) = \max_{e \geq 0} q_a \left\{ \int_0^{e\frac{q_n}{q_a} + n\frac{q_o}{q_a}} D\left(\frac{1+\delta}{\delta y}\right) dy - e(1+\beta) + \frac{\beta}{q_a} \int \widetilde{W}_0(\tilde{s}') Q(dq_n' \mid q_n, q_a) \right\}$$

shares this property. It then follows, by iteration, that it is permissible to define $W(s) \equiv \frac{\widetilde{W}(S(\tilde{s}))}{q_a}$.

This formulation is convenient because, first, the first component of the state vector (ξ, ζ) evolves stochastically, and the evolution of the second component is exactly determined by the choice of the value of z. With the problem in this form, and given the assumptions made above, it is not hard to check that the results in Stokey et. al., Section 9.2, are applicable.

innocuous; see Theorems 9.8 and 9.10 in Stokey at. al. (1989). In this case $Z(\xi,\zeta)$ is defined by the value of z solving, at each (ξ,ζ),

$$D\left(\frac{1+\delta}{\delta(\zeta+\delta z)}\right)+\beta\int W_2(\xi',z)\Xi(d\xi'\mid\xi)-\frac{1+\beta}{\xi}\begin{cases}=0 \text{ if } z>0\\ \\ \leq 0 \text{ if } z=0.\end{cases} \tag{9}$$

Differentiating (8) and using the envelope theorem,

$$W_2(\xi,\zeta)=D\left(\frac{1+\delta}{\delta(\zeta+\delta z)}\right),$$

in which case (9) becomes, for each (ξ,ζ),

$$D\left(\frac{1+\delta}{\delta(\zeta+\delta z)}\right)+\beta\int D\left(\frac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right)\Xi(d\xi'\mid\xi)-\frac{1+\beta}{\xi}\begin{cases}=0 \text{ if } z>0\\ \\ \leq 0 \text{ if } z=0.\end{cases} \tag{10}$$

From (10), the return to increasing the future installed base, z, is the sum of the first two terms: the additional present (since the future installed base is increased via current entry) and expected future surplus. Marginal cost is the final term: the value of what would have been produced elsewhere.

For (ξ,ζ) such that $Z(\xi,\zeta)>0$, the second order condition associated with the maximal z can be written

$$\begin{aligned}&-D'\left(\tfrac{1+\delta}{\delta(\zeta+\delta z)}\right)\tfrac{1+\delta}{(\zeta+\delta z)^2}\\ &-\beta\int D'\left(\tfrac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right)\tfrac{1+\delta}{\delta}\tfrac{1+\delta Z_2(\xi',z)}{[z+\delta Z(\xi',z)]^2}\Xi(d\xi'\mid\xi)\leq 0.\end{aligned} \tag{11}$$

4 Results

Results on observables such as entry into the industry all follow from the properties of the optimal future installed base, $Z(\xi, \zeta)$. Thus, we develop these properties first.

4.1 Installed base

Our first result is that the optimal future installed base is smaller when the current installed base is larger.

Proposition 1 $Z(\xi, \zeta)$ is non-increasing in ζ, and if $Z(\xi, \zeta) > 0$, $Z(\xi, \zeta)$ is decreasing in ζ.

Referring to (10), a larger installed base, ζ, influences the return and cost of increasing the future installed base, z, only by lowering the current product price; this reduces the optimal future installed base. But observe that when the installed base increases, the adjustment to the optimal future installed base, achieved by reducing current entry, cannot be large enough to cause the current price actually rise. The reasoning is simply that the optimal reduction in the future installed base spreads the implied price response over both periods. This fact will prove useful below.

Corollary 1 $\frac{\partial Z(\xi,\zeta)}{\partial \zeta} > -\frac{1}{\delta}$.

Next, how does an increase in the productivity of younger workers, ξ, influence the optimal future installed base?[7]

[7] An increase in the productivity of young workers can be given the obvious interpretation of more output, but there are other alternatives. For example, if consumers prefer

Proposition 2 If $Z(\xi, \zeta) > 0$, $Z(\xi, \zeta)$ may be increasing or decreasing in ξ, but cannot be everywhere decreasing in ξ.

Referring to (10) again, an increase in the productivity of younger workers, ξ, has two effects on the returns and costs associated with increasing the future installed base. One effect is that a given increment to the future installed base becomes less costly since it may be achieved with fewer entrants, each of whom is more productive; this effect works to increase the future installed base. Second, when the productivity of young workers is greater, if growth in know how is persistent, this is likely to lead to even greater future productivity, and possibly a much lower price in the future. Anticipating this, the optimal future installed base (comprising workers who enter in the current period, and who will be old when price is likely to be low) may actually decline. But observe that this outcome depends on the fact that if future productivity is higher, there will be an increase in the future installed base at that time; i.e., that $Z(\xi, \zeta)$ is increasing in ξ for some values of ξ.[8]

While it is possible for the future installed base to be declining in the current productivity of younger workers, there is a sense in which the persistence of growth in know how must be "very large" if it is to lead to this outcome. Let Ξ' be the conditional density associated with Ξ, and $\eta_\xi^{\Xi'}$ be the elasticity of Ξ' with respect to ξ.

---

services based on new, "cool", technology, then younger workers provide more efficiency units of output, even if their physical output is now much greater; likewise if the consumers' tastes are more Luddite in nature.

[8]Rosenberg (1976) discusses a similar idea, but applied to the firm's decision to adopt new technology.

Corollary 2 If $Z(\xi, \zeta) > 0$,

$$\frac{\partial Z(\xi,\zeta)}{\partial \xi} \gtreqless 0 \Leftrightarrow \beta \int \frac{\eta_{\xi}^{\Xi'} D\left(\frac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right) \Xi(d\xi' \mid \xi)}{D\left(\frac{1+\delta}{\delta(\zeta+\delta z)}\right) + \beta \int D\left(\frac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right) \Xi(d\xi' \mid \xi)} \gtreqless -1.$$

Intuitively, when the productivity of younger workers increases, this creates surplus both now and in the future, encouraging a greater future installed base. The opposing effect occurs only in the future, and depends on both the importance of future surplus in the overall benefit-cost calculation and the impact greater productivity has on the way future productivity evolves; e.g., if productivity growth is uncorrelated or moderately correlated over time, then the future installed base must rise with the current productivity of the young. *In what follows we focus on the case where this occurs, i.e., we assume*

$$\frac{\partial Z(\xi,\zeta)}{\partial \xi} \geq 0.$$

4.2 Price

Since $\zeta \equiv \frac{nq_o}{q_a}$ and $z \equiv \frac{e\xi}{\delta}$, the equilibrium product price is given by

$$P(\xi,\zeta) \equiv D\left(\frac{1+\delta}{\delta[\zeta + \delta Z(\xi,\zeta)]}\right).$$

From Corollary 1, it is immediate that price is declining in ζ; that is,

$$\frac{\partial}{\partial \zeta}[\zeta + \delta Z(\xi,\zeta)] > 0.$$

Thus, while an increase in the installed base reduces the future installed base, it cannot do so by enough actually to reduce current output and increase price.

The impact of an increase in the productivity of young workers, ξ, on price is opposite in sign to its impact on the future installed base. That is, if an increase in ξ results in a greater future installed base, since this increase is accomplished by greater current output, equilibrium price must fall.

4.3 Entry

Equilibrium entry by young workers is given by

$$e(\xi,\zeta) \equiv \frac{Z(\xi,\zeta)\delta}{\xi}. \tag{12}$$

From Proposition 1, it is immediate that entry declines when the installed base rises; more old workers, or more productive old workers, crowd out the young.

How does greater productivity of young workers, ξ, affect entry? That is, what is the response of entry by the young in response to more significant technological advance? From (12) it is immediate that entry increases with technology if and only if the elasticity of the future installed base with respect to ξ exceeds unity, i.e.,

$$\frac{\partial e(\xi,\zeta)}{\partial \xi} \gtreqless 0 \Leftrightarrow \eta_\xi^Z \gtreqless 1,$$

where

$$\eta_\xi^Z \equiv \frac{\xi}{Z(\xi,\zeta)} \frac{\partial Z(\xi,\zeta)}{\partial \xi}.$$

If, for example, $\eta_\xi^Z < 1$, the optimal adjustment to the future installed base in response to a more significant technological advance is moderate. In this case, simply having the same number of young workers each produce more, for example, delivers too great an increase in future installed base; achieving the optimal future installed base requires decreasing the number of young workers. The tendency towards modest adjustment of the installed base is supported by the fact that older workers will continue to produce, and exacerbated by less elastic product demand. We discuss this in more detail below.

That entry of young workers might decrease in response to technological advance may at first strike the reader as unintuitive. However, basically the same phenomenon has been modelled and documented in the literature on new product industries; see Gort and Klepper (1983) and Jovanovic and MacDonald (1994). In new product industries, the pace of cost-reducing technological change outstrips demand growth. Thus, fewer, larger, firms serve the industry as it develops, causing a "shakeout" of existing, less efficient, firms. In the present context, the less efficient are the old workers who have little to do besides producing in the industry, and who will not exit. Thus, technology change may be met by entry of comparatively few young workers who, as a group, serve a substantial part of the market, leaving the residual to the old, less efficient, workers.

4.4 Young/old income distribution

Since old workers produce q_o and earn $P(\xi, \zeta)q_o$, and the young produce q_y and earn $P(\xi, \zeta)q_y$, the relative income of young and old is simply their relative output, $\frac{q_y}{q_o}$. The implied within-industry distribution of relative output and income is given by:[9]

Relative Income	Proportion of workers
1	$\frac{n}{n+e(\xi,\zeta)}$
$\frac{q_y}{q_o}$	$\frac{e(\xi,\zeta)}{n+e(\xi,\zeta)}$

When there are more old workers in the industry, implying greater ζ, given productivity of young and old, there will be less entry. Thus, when the old are more numerous, the distribution of relative income both has a lower mean and becomes more skewed, simply because young workers, while more productive, are less numerous.

When older workers are more productive, given their number and the productivity of the young, ζ is also greater. Thus, younger workers make up a smaller proportion, and their relative income is also lower. In this case the mean relative income is lower, but the impact on skewness is ambiguous since, while younger workers make up a smaller proportion, their relative income is also lower.

Next, an increase in the productivity of younger workers, given the number and productivity of the old, increases ξ and the young/old relative income. Thus, when know how advances more quickly, each young worker makes use of the new ideas and earns relatively more than each old worker. As discussed earlier, an increase in ξ may lead either to increased or decreased entry by the young, depending on whether $\eta_\xi^Z \gtrless 1$. If entry increases, i.e., if $\eta_\xi^Z > 1$, so does the mean relative income, but the impact on skewness is ambiguous since a larger number of young workers each earn relatively

[9] As we discuss below, there are several reasons why our results are generally less applicable to income.

more. But if more rapidly advancing know how results in less entry, i.e., if $\eta_\xi^Z < 1$, then a more rapidly evolving technology implies a within-industry distribution of income that is "doubly" more skewed in the sense of a smaller proportion of those in the industry each earning relatively more. In this case, if one imagines observing a sequence of within-industry relative income distributions, a varying pace of technological advance implies a correspondingly varying degree of skewness. When $\eta_\xi^Z < 1$, the faster is the pace of change, the fewer young workers make use of this technology, and the more old workers become has-beens in the sense of both being numerous and having low relative output. This is in addition to the fact that greater ξ implies lower price, in which case the old are absolutely worse off.

When is $\eta_\xi^Z < 1$? A necessary and sufficient condition can be derived, but there is a sufficient condition which is simpler and perhaps more illuminating. To see what this condition is, let

$$\eta^P(\xi,\zeta) \equiv \frac{\frac{1+\delta}{\delta[\zeta+\delta Z(\xi,\zeta)]}}{D\left(\frac{1+\delta}{\delta[\zeta+\delta Z(\xi,\zeta)]}\right)} D'\left(\frac{1+\delta}{\delta[\zeta+\delta Z(\xi,\zeta)]}\right)$$

be the elasticity of inverse demand with respect to Y/Q, equal in magnitude, but opposite in sign, to the reciprocal of the price elasticity of demand for the product. Also, let

$$\eta_z^Q(\xi,\zeta) = \frac{\delta Z(\xi,\zeta)}{\zeta+\delta Z(\xi,\zeta)}$$

be the elasticity of current output with respect to the current installed base. Also, let

$$\eta_\zeta^Q(\xi,\zeta) = \frac{\zeta[1+\delta Z_2(\xi,\zeta)]}{\zeta+\delta Z(\xi,\zeta)}$$

be the elasticity of future output with respect to the future installed base.

Proposition 3 A sufficient condition for $\eta_\xi^Z < 1$ is

$$\frac{P(\xi,\zeta) + \beta \int P(\xi',z)\Xi(d\xi' \mid \xi)}{P(\xi,\zeta)\eta^P(\xi,\zeta)\eta_z^Q(\xi,\zeta) + \beta \int P(\xi',z)\eta^P(\xi',z)\eta_\zeta^Q(\xi',z)\Xi(d\xi' \mid \xi)} < 1.$$

In Proposition 3, the factor

$$\eta^P(\xi,\zeta)\eta_z^Q(\xi,\zeta)$$

is the elasticity of current price with with respect to Y/Q, times the elasticity of current output with respect to the number of young workers in current output, i.e., the elasticity of current price with respect to the number of young workers. Similarly, the factor

$$\eta^P(\xi',z)\eta_\zeta^Q(\xi',z)$$

is the elasticity of future price with respect to the number of old (formerly young) workers. Thus, when, for example, both elasticities exceed unity, $\eta_\xi^Z < 1$ is guaranteed.[10]

The kind of conditions leading to these elasticities being large, and thus $\eta_\xi^Z < 1$, are not hard to determine. First, a small (in absolute value) price elasticity of product demand corresponds to a large elasticity of inverse demand. That is, when demand is inelastic, price must fall more in order to absorb the increased output of younger workers, restricting the number that can rationally enter. Thus, inelastic product demand in either the current or future period contributes to $\eta_\xi^Z < 1$. And note that under some conditions

[10]The more complex condition that is both neccessary and sufficient for $\eta_\xi^Z < 1$ includes a term reflecting the impact of $\eta_\xi^{\Xi'}$, essentially requiring that the effect of greater ξ on Ξ not be too large.

– for example, linear demand – the elasticity of demand decreases as price falls. Since a more rapidly evolving technology tends to lower the future price (i.e., future entry may decline when the future installed base increases, but not, according to Corollary 1, by enough to cause future price to rise), it also tends to produce conditions where the elasticity of inverse demand is large.

Second, if there are fewer old workers in the industry when the productivity of the young increases, the increased output of the young makes up a larger proportion of industry output, and has a greater absolute effect on price, for any given elasticity. An example would be a new industry in which the old have a technology that is particularly inefficient. Third, when current entrants are expected to make up a large proportion of future output, a similar effect occurs, except with respect to the price in the future. An example would be one in which the technology available to the young is a significant breakthrough that is not likely to be repeated soon, in which case entry by the young has a large proportionate effect on future output, and hence on future price, for any given elasticity. Finally, we have interpreted demand as demand for a final product. But this is not necessary. The industry's product could equally be an input to home production or an intermediate product whose demanders are firms. Under these assumptions, familiar reasoning about the elasticity of derived demand describes situations in which the elasticity of inverse demand is likely to be large, i.e., small elasticity of demand for the final product, low elasticity of substitution between the input and others, etc.

Returning to the architecture example, note that demand for architects' services is almost exclusively for purposes of residential or commercial construction, a final product commonly cited as an example of modest price elasticity. Furthermore, the elasticity of substitution between design inputs and material inputs is plausibly thought of as small. Thus, the elasticity of derived demand for architectural services is small, and the elasticity of inverse demand large. According to the model then, if the older architects do not readily acquire the same skills as the young, these advances should have allowed comparatively few younger architects to take over much of the market for architectural services, while earning only the income of those of comparable intelligence and motivation, and causing the older generation to lose most of its income.

4.5 Discussion

4.5.1 Output versus income

Proposition 3 identifies conditions under which rapid technological advance turns those whose skills are tied to an outdated technology into has-beens in a particularly extreme fashion. Strictly, the proposition is about relative output. But given the close link between output and income in the model, it can be stated in terms of income. There are several reasons why the connection between income and output is less direct than the model supposes, and this suggests that even if the proposition is valid in terms of output, observing this in data on income may be difficult. One reason is that in industries where technological advance is substantial, and updating human capital is costly because the ideas on which old technology is based tend to differ greatly from those generating new technology, one would expect efficient contractual arrangements to ameliorate the income risk individuals face as time passes; see Weinberg (2001). Academic tenure, and the common practice of nonnegative nominal salary increments, seems an obvious example in which contract design bounds the impact of new technology on older workers.

In addition to risk sharing considerations, optimal contracting when workers' efforts are unobserved will generally cause worker productivity and pay to differ, especially in a dynamic setting; see, e.g., Lazear (1979).

4.5.2 Other choices

In the model, young/old relative output, and the relative productivity of their technologies, are defined to be identical. But just as Rosen (1981) distinguishes between ability and other factors of production and explores how choices of other factors interact to exaggerate ability differences, we may distinguish technology and other factors and investigate how these factors influence the has-been effect.[11] Suppose, for example, that technology is a neutral production function parameter, e.g.,

$$Q = qK^{\alpha},$$

where Q represents output, q represents technology level (q_y or q_o as appropriate), K is a competitively-supplied factor of production, and α is a

[11] Jovanovic (1998) studies a model in which the most up-to-date physical capital is in limited supply. When new technology and worker ability are complementary, the limited supply of leading edge capital adds to the ability-induced inequality.

parameter; $0 < \alpha < 1$. Then equilibrium relative output is given by

$$\left(\frac{q_y}{q_o}\right)^{1+\frac{\alpha}{1-\alpha}}.$$

Thus, since $1 + \frac{\alpha}{1-\alpha} > 1$, when technology evolves more quickly, i.e., ξ increases, the presence of other factors exaggerates the impact on output and income, causing, when $\eta_{\xi}^{Z} < 1$, an even more skewed distribution of relative output.

In the architect example, the young's facility with advanced technology expands the use not just of computers, but also other technology-based inputs such as PDAs, web pages, video conferencing, and so on, expanding relative productivity more than would occur purely as a result of the use of a computer as a design tool.

4.5.3 Growth in human capital

The model predicts that the income of the young will exceed the income of the old. There are cases where those who have mastered the most up to date approaches both produce more and earn more. For example, the young have turned golf into a sport, commonly beating the older generation of players whose scores have not increased, but who have been unable to match the score-reducing physical regimen taken on by the young. But more generally, even absent the above-mentioned contracting issues loosening the connection between productivity and income, there are reasons by why the old might produce and earn more despite their employing dated technology. The most obvious factor is general human capital obtained through learning by doing. When this is important, the propositions set out above will require some adjustment. For example, suppose skills increase with age at the proportional rate $\gamma > 1$, so that productivity is q_y when young and γq_y when old. Then the propositions on income distribution, for example, can be stated as a comparison of what the young earn with what the old earn, where the latter is adjusted for growth by dividing by γ; equivalently, the output of the young can be compared to what the old produced when young.

4.5.4 Retirement

The assumption that where technology evolution is important older workers typically do not choose to retool is a reasonable one. But we have also assumed that the old also do not stop working, which is more debatable, especially when the technology advance is significant. For many occupations, the option to retire mid-career if the young are greatly more productive is unlikely to be attractive; the architecture example appears to be an example. In such cases the propositions are robust to the possibility of retirement. But for others, especially highly-compensated occupations such as professional sports, the possibility of early retirement can alter our results in a manner similar to that which it alters some superstars propositions. For example, in the superstars model an rightward shift in the ability distribution increases skewness if there is no exit, but may lead to a skewness-reducing thinning of the left tail when exit is allowed for. And in the has-beens model, even when $\eta_\xi^Z < 1$, the tendency for an increase in the pace of technology change to increase skewness can be offset by the old exiting.

5 References

References

[1] Campbell, Jeffrey R., "Entry, exit, embodied technology, and business cycles," Review of economic dynamics, 1 (April 1998), 371-408.

[2] Chari, V. V. and Hopenhayn, H., "Vintage human capital, growth, and the diffusion of new technology," *Journal of Political Economy*, 99 (1991), 1142-65.

[3] Cooley, Thomas F., Greenwood, J. and Yorukoglu, M., "The replacement problem," *Journal of Monetary Economics*, 40 (December 1997), 457-499.

[4] Gort, Michael and Klepper, S., "Time paths in the diffusion of product innovations," *Economic Journal*, 92 (September 1982).

[5] Greenwood, Jeremy, and Jovanovic, B., "Accounting for growth," in *New Directions in Productivity Analysis*, ed. by Charles R. Hulten, Dean, E. R., and Harper, M. J. Chicago: University of Chicago Press, forthcoming.

[6] Jovanovic, Boyan and MacDonald, G., "Competitive diffusion," *Journal of Political Economy*, 102 (February 1994).

[7] Jovanovic, Boyan, "Vintage capital and inequality," *Review of Economic Dynamics*, 1 (April 1998), 497-530.

[8] Lazear, Edward P., "Why is there mandatory retirement?" *Journal of Political Economy*, 87 (December 1979), 1261-1284.

[9] Lazear, Edward P., *Personnel Economics,* MIT Press: Cambridge, Massachusetts, 1995.

[10] MacDonald, Glenn, "The economics of rising stars," *American Economic Review*, 78 (March 1998), 155-168.

[11] Parente, Stephen L., "Technology adoption, learning-by-doing, and economic growth," *Journal of Economic Theory*, 63 (August 1994), 346-369.

[12] Rosen, Sherwin "The economics of superstars," *American Economic Review*, 71 (December 1981), 845-858.

[13] Rosenberg, Nathan, "On technological expectations," *The Economic Journal*, 86 (September 1976), 523-535.

[14] Stokey, N., Lucas, Robert, E. Jr., and Prescott, Edward C., *Recursive Methods in Economic Dynamics,* Harvard University Press: Cambridge, Massachusetts, 1989.

[15] Weinberg, Bruce, "Long-term wage fluctuations with industry-specific human capital," *Journal of Labor Economics,* 19, (January 2001), 231-264.

[16] Zeckhauser, Richard, "Optimality in a world of progress and learning," *Review of Economic Studies*, 35 (July 1968), 363-365.

6 Proofs

6.1 Proof of Lemma 1

It suffices to show that the Kuhn-Tucker condition characterizing surplus maximizing entry is equivalent to (3). For given $\tilde{s}$, from (5), the condition is

$$q_y D\left(\frac{q_a(1+\frac{1}{\delta})}{nq_o + e(\tilde{s})q_y}\right) - q_a(1+\beta) + \beta \int \widetilde{W}_4(\tilde{s}')Q(dq_y' \mid q_y, q_a) \begin{cases} = 0 & \text{if } e(\tilde{s}) > 0 \\ \leq 0 & \text{if } e(\tilde{s}) = 0. \end{cases} \tag{13}$$

Differentiating (3) and using the envelope theorem gives

$$\widetilde{W}_4(\tilde{s}) = q_o D\left(\frac{q_a(1+\frac{1}{\delta})}{nq_o + e(\tilde{s})q_y}\right) = q_o P(\tilde{s}),$$

where the second inequality uses (4). Recalling that $q_o' = q_y$ (13) then becomes

$$q_y P(\tilde{s}) - q_a(1+\beta) + \beta \int q_y P(\tilde{s}')Q(dq_y' \mid q_y, q_a) \begin{cases} = 0 & \text{if } e(\tilde{s}) > 0 \\ \leq 0 & \text{if } e(\tilde{s}) = 0, \end{cases} \tag{14}$$

equivalent to (3).

6.2 Proof of Proposition 1

Differentiation of (10), assuming $Z(\xi, \zeta) > 0$, gives

$$\frac{\partial Z(\xi,\zeta)}{\partial \zeta} \propto \frac{-D'\left(\frac{1+\delta}{\delta(\zeta+\delta z)}\right)\frac{1+\delta}{\delta(\zeta+\delta z)^2}}{D'\left(\frac{1+\delta}{\delta(\zeta+\delta z)}\right)\frac{1+\delta}{(\zeta+\delta z)^2} + \beta \int D'\left(\frac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right)\frac{1+\delta}{\delta}\frac{1+\delta Z_2(\xi',z)}{[z+\delta Z(\xi',z)]^2}\Xi(d\xi' \mid \xi)} < 0,$$

since (11) implies the denominator is positive and $D' > 0$.

6.2.1 Proof of Corollary 1

With some rearrangement

$$\frac{\partial Z(\xi,\zeta)}{\partial \zeta} = \frac{-1}{\delta + \frac{\int D'\left(\frac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right)\frac{1+\delta Z_2(\xi',z)}{[z+\delta Z(\xi',z)]^2}\Xi(d\xi'|\xi)}{D'\left(\frac{1+\delta}{\delta(\zeta+\delta z)}\right)\frac{1}{(\zeta+\delta z)^2}}},$$

where the concavity of W as a function of ζ implies that the fraction in the denominator is positive. Thus,

$$\frac{\partial Z(\xi,\zeta)}{\partial \zeta} > -\frac{1}{\delta}.$$

6.3 Proof of Proposition 2

Differentiation of (10), assuming $Z(\xi,\zeta) > 0$, gives

$$\frac{\partial Z(\xi,\zeta)}{\partial \xi} = \frac{\frac{1+\beta}{\xi^2} + \int D\left(\frac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right)\frac{\partial}{\partial \xi}\Xi(d\xi' \mid \xi)}{D'\left(\frac{1+\delta}{\delta(\zeta+\delta z)}\right)\frac{1+\delta}{(\zeta+\delta z)^2} + \beta\int D'\left(\frac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right)\frac{1+\delta}{\delta}\frac{1+\delta Z_2(\xi',z)}{[z+\delta Z(\xi',z)]^2}\Xi(d\xi' \mid \xi)}, \tag{15}$$

where (11) implies the denominator is positive. Now, as a function of ξ,

$$\frac{\partial}{\partial \xi} D\left(\frac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right) = D'\left(\frac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right)\frac{1+\delta}{\delta}\frac{-\delta\frac{\partial Z(\xi',z)}{\partial \xi}}{[z+\delta Z(\xi',z)]^2}.$$

Suppose $\frac{\partial Z(\xi,\zeta)}{\partial \xi} \leq 0$ everywhere. Then

$$D\left(\frac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right)$$

is an increasing function of ξ, in which case

$$\int D\left(\frac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right)\frac{\partial}{\partial \xi}\Xi(d\xi' \mid \xi)$$

is also increasing in ξ. This yields $\frac{\partial Z(\xi,\zeta)}{\partial \xi} > 0$, a contradiction. Thus, while $\frac{\partial Z(\xi,\zeta)}{\partial \xi} < 0$ is possible, it cannot hold everywhere.

6.3.1 Proof of Corollary 2

From (10), when $Z(\xi, \zeta) > 0$,

$$\frac{1+\beta}{\xi} = D\left(\frac{1+\delta}{\delta(\zeta+\delta z)}\right) + \beta \int D\left(\frac{1+\delta}{\delta[z+\delta Z(\xi', z)]}\right) \Xi(d\xi' \mid \xi).$$

Making this substitution in (15), and multiplying through by ξ, the numerator of $\frac{\partial Z(\xi,\zeta)}{\partial \xi}$ is

$$D\left(\tfrac{1+\delta}{\delta(\zeta+\delta z)}\right) + \beta \textstyle\int D\left(\tfrac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right) \Xi(d\xi' \mid \xi)$$

$$+ \textstyle\int D\left(\tfrac{1+\delta}{\delta[z+\delta Z(\xi',z)]}\right) \Xi'(\xi' \mid \xi) \tfrac{\xi}{\Xi'(\xi' \mid \xi)} \tfrac{\partial}{\partial \xi} \Xi'(\xi' \mid \xi) d\xi'.$$

Since

$$\eta_\xi^{\Xi'} \equiv \frac{\xi}{\Xi'(\xi' \mid \xi)} \frac{\partial}{\partial \xi} \Xi'(\xi' \mid \xi),$$

this expression can be written

$$D\left(\frac{1+\delta}{\delta(\zeta+\delta z)}\right) + \beta \int D\left(\frac{1+\delta}{\delta[z+\delta Z(\xi', z)]}\right) \Xi(d\xi' \mid \xi) + \int \eta_\xi^{\Xi'} D\left(\frac{1+\delta}{\delta[z+\delta Z(\xi', z)]}\right) \Xi(d\xi' \mid \xi), \tag{16}$$

from which the result follows by rearrangement.

6.4 Proof of Proposition 3

As in the proof of Corollary 2, the numerator of $\frac{\partial Z(\xi,\zeta)}{\partial \xi}$ is given by (16). Since the concavity of W in ζ implies the final integral in (16) is negative, the numerator is less than

$$D\left(\frac{1+\delta}{\delta(\zeta+\delta z)}\right) + \beta \int D\left(\frac{1+\delta}{\delta[z+\delta Z(\xi', z)]}\right) \Xi(d\xi' \mid \xi).$$

The denominator is

$$D'\left(\frac{1+\delta}{\delta(\zeta+\delta z)}\right) \frac{1+\delta}{(\zeta+\delta z)^2} + \beta \int D'\left(\frac{1+\delta}{\delta[z+\delta Z(\xi', z)]}\right) \frac{1+\delta}{\delta} \frac{1+\delta Z_2(\xi', z)}{[z+\delta Z(\xi', z)]^2} \Xi(d\xi' \mid \xi).$$

Substitution of the definitions of $\eta^P(\xi,\zeta), \eta_z^Q(\xi,\zeta)$ and $\eta_\zeta^Q(\xi,\zeta)$ yields the result.