Knowledge Spillovers and Inequality

By Jan Eeckhout and Boyan Jovanovic*

We develop a dynamic model with knowledge spillovers in production. The model contains two opposing forces. Imitation of other firms helps followers catch up with leaders, but the prospect of doing so makes followers want to free ride. The second force dominates and creates permanent inequality. We show that the greater are the average spillovers and the easier they are to obtain, the greater is the free-riding and inequality. More directed copying raises inequality by raising the free-riding advantages of hanging back. Using Compustat and patent-citation data we find that copying is highly undirected. (*JEL* D33, I11, O31)

Spillovers of knowledge flow from leaders to followers. Discoveries at the frontier of science and technology certainly increase output produced by the frontier firms, but followers, too, benefit from those inventions and the technological improvements they generate. The benefits of a frontier firm's discoveries are often hard to protect from copying by competing users. By the very nature of knowledge, property rights are not easily enforced. Patent legislation provides only an imperfect substitute for those property rights, and most often grants those rights for a limited time only. It is generally well understood that safeguarding the discoveries from copying is needed to create the right incentives for investment into pushing out the frontier of knowledge. A frontier firm that does not reap all the benefits from its investment efforts will choose investment levels that are less than optimal. Most writers on the subject conclude that knowledge spillovers induce free-riding behavior and that this lowers growth below what is optimal.

Often equally noncontroversial is the idea that knowledge spillovers are a force of convergence toward equality. Rent-seeking followers may access frontier knowledge without making the costly investment. With access to the newest technology, followers will eventually catch up with the leading firms. This then leads to the conclusion that spillovers are an equalizing force.

In this paper we challenge the notion that knowledge spillovers promote equality. Rather, by inducing followers to free ride, they promote *inequality*. This force was probably at work during two epochs of rising interfirm inequality. The first epoch is the period 1850–1930 when the development of the telegraph and telephone coincided with a substantial increase in the size of the largest firms. And with the advent of the information age, the last 20 years of the twentieth century have seen a downsizing at the bottom end of the size distribution, and a rise in dispersion of stock-market valuations (Jovanovic and Peter L. Rousseau, 2000a).

We propose a theory with the central premise that copying technology from leaders is *imperfect*. While ideally follower firms would like to copy the frontier knowledge, in many cases they will settle for a technology that is better than their own, but inside the frontier. There are many reasons why a firm does not, or cannot copy the frontier technology. The firm must find a technology that fits its own needs. It takes some time to collect knowledge about existing technologies and its characteristics. Once a suitable technology is discovered, information must be gathered to establish whether or not the copied technology is protected by patents. We will model this costly copying process by means of a random arrival of technologies.

* Eeckhout: Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104 (e-mail: eeckhout@ssc.upenn.edu; ⟨http://www.ssc.upenn.edu/~eeckhout⟩); Jovanovic: Department of Economics, New York University, 269 Mercer Street, New York, NY 10003 (e-mail: bj2@nyu.edu; ⟨http://www.nyu.edu/econ/user/jovanovi⟩). We thank Fernando Alvarez, Robert Axtell, Bart Hobijn, Peter L. Rousseau, and two referees for helpful comments and Ricardo Caballero for the citations data.

A second key assumption is that technological leaders have less to learn from others than followers do. In other words, we assume that the more a firm knows, the less knowledge it can access from others. We show that this decreasing access-of-knowledge function generates inequality that would otherwise not exist. This occurs through the mechanism of investment, in which a forward-looking firm takes into account the amount of knowledge that it will be able to copy in the future. Investing more now would reduce a follower's ability to copy others in the future, and this prevents it from investing as fast as it would otherwise.

Our main result is that a long-run equilibrium must entail a nondegenerate distribution of firm sizes. All grow at the same rate, and followers have no incentive to catch up. In fact, in a steady-state equilibrium, the distribution of firms must be such that the elasticity of access with respect to the firm's knowledge is constant. An access technology with a higher elasticity (constant across firms), will provide stronger incentives for hanging back and the distribution of firms will be more unequal. In the context of our model, it also follows that growth will be lower as free-riding and inequality increase.

Another thing that raises the incentive to free ride and, thereby, raises inequality, is the ability to direct one's copying toward the leaders. If *any* firm can easily copy the leaders, we would expect to see huge incentives for firms to hang back. As the degree of directedness approaches the frictionless case, most firms hang back, and just a few leaders remain.

The unit analysis is the firm. In the Compustat sample data we find evidence in support of the model. We make use of the firms' book value (i.e., the investment in tangible physical capital) as opposed to the market value, to extract information on the intangible value of the spillovers. There is strong evidence that spillovers are higher for smaller firms, and this confirms the finding of others that small firms and plants are more productive.

Linking inequality to free-riding incentives for firms or households is obviously not new. The provision of public goods (Mark Bagnoli and Barton Lipman, 1989) and investment in a new technology (Jovanovic and Saul Lach, 1989) are situations in which some agents will take the lead while others will free ride. Richard R. Nelson (1988) shows that innovators and imitators can coexist in the long run, and Jovanovic and Glenn M. MacDonald (1994) derive conditions under which imitators will not catch up. Another mechanism through which externalities raise inequality is segregation. Charles Tiebout's (1956) theory of local public goods suggests that when people are perfectly mobile, differences in endowments can become larger. Robert Tamura (1991), Gerhard Glomm and B. Ravikumar (1992), Roland Bénabou (1996), and Oded Galor and Daniel Tsiddon (1997) have rightly stressed these forces in dynamic models of growth. We add to this literature in several ways. We specify a simple spillover technology that leads to a lower bound for how much inequality there will be. We estimate the parameters with stock-market data and with patent citations. We then show that undirected copying also explains the rather long patent-citation lags that Ricardo Caballero and Adam Jaffe (1993) document.

We can also think of this model in terms of the size of a given city. Zipf's Law states that the distribution of city size is distributed according to the Pareto density with exponent -1. Our model provides an explanation for the distribution of cities—say, within the United States—based upon externalities among cities. Recently, Yannis M. Ioannides and Henry G. Overman (2000) and Xavier Gabaix (1999) propose an explanation postulating that cities grow at a rate proportional to their size. This growth process is confirmed empirically. Our theory can provide economic fundamentals based on differential externalities that generate such a growth process and in which the equilibrium distribution must be Pareto with unit exponent.

In the next section, we propose an endogenous growth model that is standard except for a more general definition of spillovers. In Section II, we define the costly imitation process and derive the access-to-technology function. Section III defines and characterizes equilibrium, including a comparison to some benchmark cases and the derivation of the equilibrium distribution of firms. In Section IV, we investigate the impact on inequality of growth, directed copying, and aggregate spillovers. In Section V, we use stock-market data to identify the spillover mechanism, and we provide some other tests of the model. Section VI concludes.

I. Model

We study a one-sector growth model with convex capital-adjustment costs at the firm level. The technical innovation is that the entire distribution of capital—and not just its average or its maximum—enters the production function "externally." This plausible generalization of the spillover mechanism has some dramatic implications for inequality in the long run.

Production.—Firms produce output, y, using an aggregate of physical and human capital, k. The production function is:

$$y = A_t k \tag{1}$$

where A_t is a productivity parameter that changes over time so as to reflect the knowledge of others and the portion of it that the firm in question can access. We shall describe the spillover process in greater detail below. For now, we shall summarize it by the function

$$A_t = A_t(k)$$

for which $A_t'(k) < 0$, which reflects an advantage low-k firms have in accessing productively useful knowledge.

Investment.—The firm faces internal costs of rapid adjustment of k that are proportional to its output. Starting from k this period, the firm can have $\tilde{k}$ units of capital next period at the cost of $yC(\tilde{k}/k)$, so that its net output would be

$$\left[1 - C\left(\frac{\tilde{k}}{k}\right)\right]y.$$

We assume that $C' > 0$ and $C'' > 0$, and that if the firm does not invest, only a fraction $(1 - \delta)$ of its capital will remain into the next period, so that $C(1 - \delta) = 0$. A firm is too small to affect the price of the product. Its short-run production cost is zero up to the point y, and there it becomes infinite. Its short-run supply curve is therefore vertical at that point. Thus total revenue net of adjustment costs is the firm's profit. A larger k means a larger profit, the prospect of which induces firms to invest in k.

The Decision Problem of the Firm.—Suppose that the rate of interest, r, is constant, and let $v_t(k)$ be the present value of profits at date t, so that

$$v_t(k) = \max_{\tilde{k}} \left\{ \left[1 - C\left(\frac{\tilde{k}}{k}\right)\right] A_t(k)k + \frac{1}{1+r} v_{t+1}(\tilde{k}) \right\}. \tag{2}$$

The first-order condition for a maximum is that the marginal cost of a unit of investment should equal its marginal return:

$$C'\left(\frac{\tilde{k}}{k}\right) A_t(k) = \frac{1}{1+r} v'_{t+1}(\tilde{k}). \tag{3}$$

The envelope theorem says that

$$v_t'(k) = A_t(k)\left(\frac{\tilde{k}}{k} C' + (1 - C)[1 - \varepsilon_t(k)]\right) \tag{4}$$

where

$$\varepsilon_t(k) = \left| \frac{kA_t'(k)}{A_t(k)} \right| \tag{5}$$

is the absolute value of the elasticity of A_t with respect to k. Eliminating $v_{t+1}(\cdot)$ from (3), yields a second-order difference equation for k:

$$(1 + r)C'\left(\frac{k_{t+1}}{k_t}\right) A_t(k_t) = A_{t+1}(k_{t+1}) \left\{ \frac{k_{t+2}}{k_{t+1}} C'\left(\frac{k_{t+2}}{k_{t+1}}\right) + \left[1 - C\left(\frac{k_{t+2}}{k_{t+1}}\right)\right] [1 - \varepsilon_{t+1}(k_{t+1})] \right\}. \tag{6}$$

Let x_k and x_A denote the growth factor of k and A respectively. Suppose that

$$\varepsilon_t(k) = \varepsilon \geq 0 \qquad \text{for all } k \text{ and } t \tag{7}$$

and suppose that the growth factors $k_{t+1}/k_t \equiv x_k$

and $A_{t+1}(k_{t+1})/A_t(k_t) \equiv x_A$ were constant and equal across firms. Then (6) would imply that

$$(8)\quad 1 - \varepsilon = \left[\frac{1}{x_A}(1 + r) - x_k\right]\frac{C'(x_k)}{1 - C(x_k)}.$$

Equation (8) has four unknowns: x_k, x_A, ε, and r. We shall derive r from the optimal saving conditions, and x_A and ε from the spillover process which we have yet to specify. In passing, note from (4) that ε lowers the return to investment. The larger is ε, the more spillovers the firm loses when it raises its own k.

Optimal Saving Behavior.—The representative consumer's lifetime utility is

$$\sum_{t=0}^{\infty} \rho^t \frac{c_t^{1-\gamma}}{1-\gamma} \qquad \rho < 1 \quad \gamma > 0.$$

He maximizes his preferences under the constraint that $\sum_0^\infty [1/(1 + r)^t]c_t$ not exceed wealth. If the consumption-growth factor is constant at x_c, the first-order conditions imply that $1 + r = x_c^\gamma/\rho$. But consumption must grow at the same rate as potential output, which means that

$$x_c = x_A x_k$$

so that

$$(9)\qquad 1 + r = \frac{(x_A x_k)^\gamma}{\rho}$$

which we can use to eliminate r in (8) and get

$$(10)\quad 1 - \varepsilon = \left[\frac{1}{\rho}(x_A x_k)^{\gamma-1} - 1\right]\frac{x_k C'(x_k)}{1 - C(x_k)}.$$

II. Access to the Knowledge of Other Firms

We now specify the copying technology, and how it determines the function $A_t(k)$, and, hence, how it determines ε and x_A. Let K denote the supremum of the distribution of k among firms. It represents the technological frontier. Every other firm would like to access K, but patents and the frictions of search may prevent it from doing so. A firm may, instead, have to settle for accessing the knowledge of some firm whose k is below K. In general, what the firm will end up accessing will depend on what the firm itself knows to begin with, k, and on what its competitors know—the distribution of k in the population of firms.

Let $z = k/K$, and let $H(\cdot)$ be the cumulative distribution of z among firms. Together with K_t, $H_t(\cdot)$ fully describes the menu of technologies that the firm can copy. Then we shall assume that

$$(11)\quad A_t(k) = \left[K_t\left(1 + \int_{k/K_t}^{1} \alpha(z)h_t(z)\,dz\right)\right]^\beta$$

for all $k \le K$.[1] The parameter $\beta > 0$ measures the intensity, economywide, of the externality, and $h_t(z)$ is the density of z. The average height of the function $\alpha(z) \ge 0$ tells us how much advantage in copying a backward firm enjoys compared to the leaders, and the slope $\alpha'(z)$ describes the direction of copying—if $\alpha'(z) > 0$, copying is directed toward the leaders, whereas, if $\alpha'(z) < 0$, it is directed toward the followers.

This deterministic formulation for $A(k)$ states that the firm copies a weighted average of the knowledge of other firms. We lean heavily on the law of large numbers here that, in fact, does not obtain to this extent. The formulation aims to capture a random, imperfectly directed process of search in which a firm draws randomly from the entire distribution of firms and their technologies. For firm k, it makes sense only to copy technologies that are better than its own technology k. As a result, $A(k)$ is the expectation of what the firm would get from a draw from the truncated distribution (above k) of other firms. Note that A_t is decreasing in k: for the frontier firm for example, there is no better technology

[1] Characterizing equilibrium requires $A_t(k)$ to be well defined for any feasible $k \in \mathbb{R}_+$, including k outside the support of H. See the Appendix for the exhaustive definition of $A(k)$. Note also that this formulation implies that A is bounded from below by 1.

to be copied.[2] The function $\alpha(z)$ is central, and later on we shall study it in detail and then estimate it using stock-market data.

Investment in k now plays a dual role. It raises the internal component of the productive input, but it also reduces the component "borrowed" from others, because it raises k/K—the truncation point in the distribution from which copied technologies are drawn. Recognizing this, the firm invests less than it would if the spillovers were absent and β were zero.

III. Equilibrium

We now define a steady-state equilibrium in which every firm's capital expands at the growth factor x_k so that H_t (the distribution of k/K_t) is fixed at H. This is possible only if $\varepsilon_t(k)$ is a constant as specified in (7). This is the restriction that delivers the implications for steady-state inequality.

In equilibrium, all firms grow at the same rate, and therefore, each grows as fast as the frontier, K. Then, from the definition of $A(k)$, the factor by which the accessed knowledge of each firm grows is

$$x_A = x_k^{\beta}.$$

Eliminating x_A from (10) yields

$$(12)\quad 1 - \varepsilon = \left[\frac{1}{\rho} x_k^{(1+\beta)(\gamma-1)} - 1\right] \frac{x_k C'(x_k)}{1 - C(x_k)}.$$

We are down to two unknowns: x_k and ε. Letting $x_k = x$, we can write (12) as

$$(13)\quad 1 - \varepsilon = \left[\frac{1}{\rho} x^{(1+\beta)(\gamma-1)} - 1\right] \frac{x C'(x)}{1 - C(x)}.$$

Now we can define equilibrium formally:

Definition 1: Equilibrium consists of three scalars: $x > 1 - \delta$, $z_{\mathrm{m}} \in [0, 1]$ and $\varepsilon \in [0, 1)$, and density function $h(z)$ on the interval [0, 1], such that (13), (19), and (20) hold.

Here, z_{m} is the lower bound on the support of the distribution. Note that it follows from the definition of $A(k)$, that the density function h is well defined for all $k \in \mathbb{R}_+$ and hence for all $z = k/K$. Of course, $h(z) = 0$ for all $z \in [0, z_{\mathrm{m}})$. The restriction that $\varepsilon < 1$ prevents the marginal product of k from being negative. The restriction makes sense if there is free disposal of k.

A. *Steady-State Growth*

Write (13) as

$$(14)\quad 1 - \varepsilon = \Psi(x)$$

where

$$(15)\quad \Psi(x) = \left[\frac{1}{\rho} x^{(1+\beta)(\gamma-1)} - 1\right] \frac{x C'(x)}{1 - C(x)}.$$

Note that the lowest possible growth factor is $x = 1 - \delta$. Since ε is nonnegative, equilibrium exists only if $\Psi(1 - \delta)$ is less than unity, which requires that

$$(16)\quad \frac{(1-\delta)C'(1-\delta)}{1 - C(1-\delta)} < 1.$$

LEMMA 1: *For $\gamma \geq 1$, Ψ is strictly increasing.*

PROOF:
See the Appendix.

The implication of this lemma is that provided condition (16) holds, for any given $\varepsilon \in [\Psi(1 - \delta), 1)$, equilibrium, x^*, exists and is unique. This is illustrated in Figure 1, which plots Ψ as a function of x.

B. *Gibrat's Law and Efficiency*

1. ***Gibrat's Law.***—We say that Gibrat's Law holds if any initial distribution of firms replicates itself. Consider two benchmark cases. Both of them imply Gibrat's Law.

Case (i): No Externality. This is the property of the standard Ak growth model. In the

[2] Proposition 5 of Jovanovic and MacDonald (1994) derives a similar result in an explicit imitation process in which search effort is endogenous.

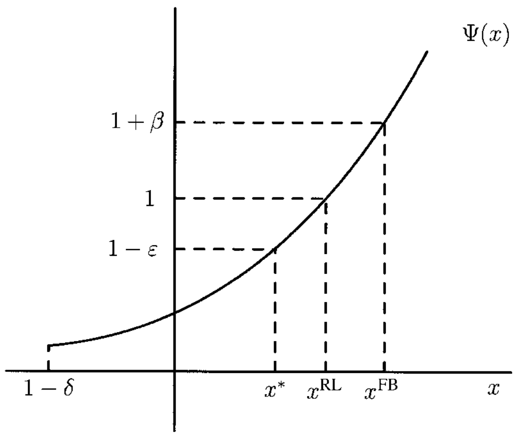

FIGURE 1. GROWTH AND EFFICIENCY

model here, if the spillover parameter β is zero, there is no external effect, and output is produced with a constant-returns-to-scale technology: $y = k$. The first-order condition reduces to

$$1 = \Psi(x).$$

Given convexity of the cost function, the equilibrium growth rate x is unique. Moreover, any distribution $H(z)$ can be sustained in equilibrium. Whatever initial distribution the steady state starts off with, that distribution will remain unchanged. All firms grow at a constant rate x, but they can have different levels of capital stock k.

Case (ii): Type-independent Externality (Romer-Lucas). Let $\beta > 0$, but let spillovers satisfy $dA/dk = 0$, and hence $\varepsilon = 0$. This is the case, for example, when A is equal to the mean of the distribution H, or to the max of its support, K. Again, the growth rate is unique, satisfying: $1 = \Psi(x)$, and any distribution can be sustained in equilibrium.

2. ***Efficiency.***—If the planner could internalize the externalities (e.g., via first-best patent regulation), then a firm's objective would contain $k^{1+\beta}$. The *first-best* level of investment involves a growth factor x^{FB} that internalizes the externality. The equilibrium condition for the first-best economy satisfies $1 + \beta = \Psi(x^{FB})$, as shown in Figure 1. Since the conditions of the first welfare theorem are not satisfied in our decentralized economy (i.e., the presence of externalities), we do not expect the equilibrium to be efficient.

Any equilibrium with copying entails a growth rate x^* that is even lower than the growth rate in a model with a Romer-Lucas (type-independent) externality. To see this, note that with type-independent externalities, $\varepsilon = 0$ because $A' = 0$. As a result,

$$x^{FB} > x^{RL} = \Psi^{-1}(1) > x^*.$$

C. *The Equilibrium Distribution of Firms*

Since a stationary equilibrium requires ε not to depend on k or t, solving the differential equation (5) implies that $A_t(k) = \lambda_t k^{-\varepsilon}$ for some sequence of constants of integration $\{\lambda_t\}$. From (11), this implies that

$$\text{(17)} \qquad \lambda_t k^{-\varepsilon} = \left[K_t \left(1 + \int_{k/K_t}^{1} \alpha(z) h_t(z)\, dz \right) \right]^{\beta}.$$

Evaluating both sides at $k = K_t$ implies that $\lambda_t = K_t^{\beta+\varepsilon}$. Eliminating λ_t and simplifying leaves

$$\text{(18)} \qquad z^{-\varepsilon/\beta} = 1 + \int_{z}^{1} \alpha(s) h_t(s)\, ds$$

for all z in the support of H. Differentiating both sides with respect to z yields

$$\text{(19)} \qquad h(z) = \frac{\varepsilon z^{-1-\varepsilon/\beta}}{\beta \alpha(z)}$$

for z in the support of H. The largest number in the support of z is 1. To calculate its smallest number, denoted by z_m, we make use of the fact that

$$\text{(20)} \qquad \int_{z_m}^{1} h(z)\, dz = 1.$$

Substituting from (19) into (20) leads to an implicit function in two unknown variables ε and z_{m}:

$$\frac{\varepsilon}{\beta}\int_{z_{\mathrm{m}}}^{1}\frac{z^{-1-\varepsilon/\beta}}{\alpha(z)}\,dz = 1. \tag{21}$$

IV. Inequality

This section describes several implications of the model and the next section reports some tests.

A. *Growth and Inequality*

In this subsection we shall show that inequality rises when the overall externality, as measured by β, is stronger that inequality and growth both rise when ε falls as we move across equilibria, and we shall derive a lower bound on the extent of inequality.

The equilibrium $H(z)$ in (19) depends on β and ε only through their ratio, and so let us see how inequality in z depends on the ratio β/ε. Define by z^p the value of z that corresponds to percentile p in the distribution H. Then z^p is the solution to $p = H(z)$ or, equivalently, the solution for z satisfies

$$1 - p = \int_{z}^{1} h(s)\,ds \tag{22}$$

$$= \frac{\varepsilon}{\beta}\int_{z}^{1}\frac{s^{-1-\varepsilon/\beta}}{\alpha(s)}\,ds$$

where the second equality follows from (19). Note that $z^0 = z_{\mathrm{m}}$. Write the right-hand side as

$$\Phi\left(\frac{\varepsilon}{\beta}, z\right) = 1 - p.$$

Since the integrand $s^{-1-\varepsilon/\beta}$ is increasing in ε/β because $s < 1$, Φ is increasing in ε/β (i.e., $\Phi_1 > 0$), and it is decreasing in z, (i.e., $\Phi_2 < 0$). Therefore

$$\frac{\partial z^p}{\partial(\varepsilon/\beta)} = -\frac{\Phi_1}{\Phi_2} > 0 \qquad \text{for all } p \in [0, 1).$$

This leads to several results.

PROPOSITION 1: *(Aggregate Spillovers Raise Inequality.) As aggregate spillovers, β, rise, every percentile in the distribution of z (e.g., the median) is farther behind the leaders. That is, $\partial z^p/\partial\beta < 0$. Moreover, the support of the distribution $h(z)$ becomes larger: letting $p = 0$, we get $\partial z_{\mathrm{m}}/\partial\beta < 0$.*

As the external effect becomes more important in production (e.g., due to worse property rights protection), the equilibrium distribution of firms becomes more spread out with a larger support. From (22), a related result emerges:

PROPOSITION 2: *Fix ε and let $\alpha(z) < \infty$. As spillovers go to zero, equality must prevail:*

$$\lim_{\beta\to 0} z_{\mathrm{m}} = 1.$$

From the discussion of Gibrat's Law, we observe that there is an *explosion* in the equilibrium set at the point $\beta = 0$. For β positive but arbitrary small, $h(\cdot)$ collapses to a degenerate distribution with a unique mass point at $z = 1$. But when β is *identically* zero, *any* distribution is an equilibrium.

The next result concerns ε. Since a higher ε raises z^p for any p, and since it also implies a higher x (see Figure 1), it follows that the high-growth equilibria entail more inequality.

PROPOSITION 3: *(Inequality and Growth Are Positively Related.) As ε rises and growth, x, falls, inequality falls: $\partial z^p/\partial\varepsilon > 0$. In particular, the support of the distribution $h(z)$ shrinks: $\partial z_{\mathrm{m}}/\partial\varepsilon > 0$.*

Now from equation (16), we know that the range ε is bounded. This leads us to the following lower bound on the extent of inequality.

PROPOSITION 4:

$$z_{\mathrm{m}} \leq \exp\left(-\frac{\beta\alpha_{\min}}{1 - \Psi(1-\delta)}\right).$$

PROOF:

From (19) and (20),

$$1 = \int_{z_m}^{1} h(z)\, dz$$

$$= \frac{\varepsilon}{\beta} \int_{z_m}^{1} \frac{1}{\alpha(z)} z^{-1-\varepsilon/\beta}\, dz$$

$$\geq \frac{\varepsilon}{\beta \alpha_{\min}} \int_{z_m}^{1} z^{-1}\, dz$$

$$= -\frac{\varepsilon}{\beta \alpha_{\min}} \ln z_m$$

where $\alpha_{\min} = \min_{z \in [0,1]} \alpha(z)$, and

$$z_m \leq \exp\left(-\frac{\beta \alpha_{\min}}{\varepsilon}\right)$$

$$\leq \exp\left(-\frac{\beta \alpha_{\min}}{1 - \Psi(1-\delta)}\right)$$

because by Figure 1 and equation (16), $\varepsilon \leq 1 - \Psi(1-\delta)$.

Example 1: $\alpha(z) = \alpha$ (undirected copying).

We illustrate these propositions with the case where $\alpha(z) = \alpha$, a constant. The larger is α, the larger is the advantage in copying that a backward firm enjoys compared to the leaders. In this sense, α is an index of the incentive to free ride. Then (19) implies that

$$h(z) = \frac{\varepsilon}{\alpha\beta} z^{-1-\varepsilon/\beta} \qquad \text{for } z \in [z_m, 1]$$

where, solving

$$\int_{z_m}^{1} \frac{\varepsilon}{\alpha\beta} s^{-1-\varepsilon/\beta}\, ds = 1$$

for z_m yields $z_m^{-\varepsilon/\beta} = 1 + \alpha$, or simply

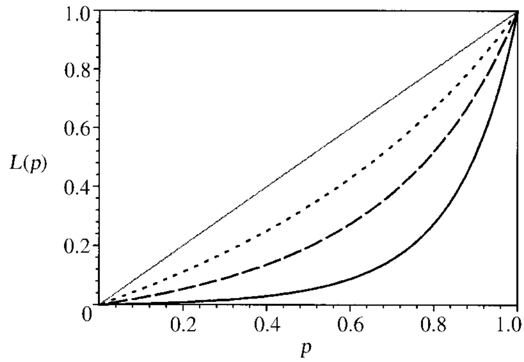

FIGURE 2. THE LORENZ CURVE FOR $\varepsilon = 0.05, 0.1, 0.2$ (WITH $\alpha = 1$, $\beta = 0.4$)

$$z_m = \left(\frac{1}{1+\alpha}\right)^{\beta/\varepsilon}. \tag{23}$$

Inequality is high when z_m is low, and disappears as $z_m \to 1$. Therefore inequality is:

(i) increasing in the free-riding incentive α, but disappears if copying is equally accessible to all ($\alpha = 0$);
(ii) increasing in the externality β (Proposition 1) but disappears if $\beta = 0$ (Proposition 2);
(iii) decreasing in ε (Proposition 3 and upcoming Lorenz curve plot) and reaches its lower bound when ε is at its largest possible level, $1 - \Psi(1-\delta)$ (Proposition 4).

The relation between ε and inequality is best illustrated plotting the Lorenz curves for some particular parameter values. For $p \in [0, 1]$, the Lorenz curve $L(p)$ is given by

$$L(p) = \frac{GL(p)}{GL(1)}$$

where

$$GL(p) = \int_{z_m}^{H^{-1}(p)} z\, dH(z)$$

$$= \frac{\varepsilon}{\alpha(\beta-\varepsilon)} \left[\left(\frac{1}{1+\alpha(1-p)}\right)^{(\beta-\varepsilon)/\varepsilon} - \left(\frac{1}{1+\alpha}\right)^{(\beta-\varepsilon)/\varepsilon} \right]$$

is the total amount of z accounted for by the least efficient p percent of the firms. In Figure 2 we choose $\alpha = 1$ and $\beta = 0.4$ [the latter based on an estimate of Robert E. Lucas, Jr. (1988) on which we shall say more below] and then plot the Lorenz curve for $\varepsilon = 0.05$ (solid line), $\varepsilon = 0.1$ (dashed), and $\varepsilon = 0.2$ (dotted). Inequality is unambiguously lower for higher ε: the Lorenz curves never cross and are always higher for higher ε.

Next, we shall show that the ability to direct copying toward the leaders raises the incentive to free ride and raises inequality.

B. *Directed Copying Raises Inequality*

Although our analysis remains deterministic, we shall now motivate directed copying by the well-known example due to George J. Stigler (1961). Suppose that copying proceeds by first taking n independent draws $s_1, \ldots, s_n$ from the distribution $H(\cdot)$, and then picking the maximal value drawn, $s^n_{\max}$. The amount copied is $\max_i(s_i)$, and its distribution is $[H(z)]^n$. The parameter n measures the directedness of search. Firm z expects to copy the amount

$$E\{\max_i(s_i)\} = \int_z^1 s d[H(s)]^n$$

$$= n \int_z^1 s[H(s)]^{n-1} h(s)\, ds.$$

This search-theoretic interpretation applies to $\alpha(z)$ in (11) if we interpret A as related to the *expected* amount of knowledge copied, and if we set

$$\alpha(z) = nz[H(z)]^{n-1}. \tag{24}$$

Then (19) gives as the equilibrium distribution

$$h(z) = \frac{\varepsilon z^{-1-(\varepsilon/\beta)}}{\beta n z[H(z)]^{n-1}}$$

or

$$\frac{d}{dz}[H(z)]^n = \left(\frac{\varepsilon}{\beta}\right) z^{-2-(\varepsilon/\beta)}.$$

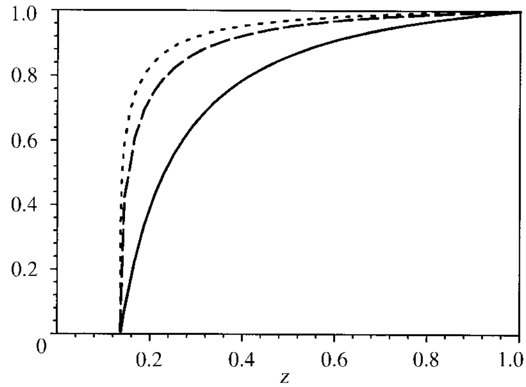

FIGURE 3. THE CUMULATIVE DISTRIBUTION FUNCTION $H(z;\, n)$ FOR $n = 1, 3, 5$ (PARAMETER VALUES $\varepsilon = 0.056$, $\beta = 0.4$)

Integrating both sides and imposing the corner condition $H(1) = 1$ yields the equilibrium cumulative distribution function (c.d.f.)

$$H(z;\, n) = \left[1 + \frac{\varepsilon}{\beta + \varepsilon}\left(1 - z^{-1-(\varepsilon/\beta)}\right)\right]^{1/n}$$

$$\text{for } z \in \left[\left(\frac{\varepsilon}{\beta + 2\varepsilon}\right)^{\beta/(\beta+\varepsilon)}, 1\right].$$

Inequality rises with n, as illustrated in Figure 3, which plots $H(z;\, n)$ for different n (solid $n = 1$, dashed $n = 3$, dotted $n = 5$). For these plots, we use the parameter values $\varepsilon = 0.056$ and $\beta = 0.4$ as estimated in the empirical analysis (see Section V).

It is immediately apparent from the plots that the cumulative density functions do not cross, and as a result, we can unambiguously rank the degree of inequality. The c.d.f. that is strictly below $H(z;\, 1)$, is more equal than $H(z;\, 3)$ which in turn is more equal than $H(z;\, 5)$. Below we show that any distribution $H(z;\, n)$ first-order stochastically dominates $H(z;\, n')$ when $n < n'$. Interestingly, however, z_{m} does not depend on n, so that the support of H does not depend on n, only the shape of H does. We have the following general result:

PROPOSITION 5: *As copying gets more directed (i.e., higher n), every percentile in the distribution of z (e.g., the median) is farther* *behind the leaders. In this sense inequality grows.*

PROOF:

$H(z; n)$ is continuous and differentiable, and

$$\frac{dH(z;\, n)}{dn} = H(z)\ln[(H(z))^n]\,\frac{-1}{n^2} > 0 \quad \forall\, z \in [0,\, 1)$$

because $\ln H < 0$ for all $z < 1$. Therefore a higher n leads to a stochastically lower distribution. That is, for each $p \in (0, 1)$, the pth percentile of the distribution, z^p, is decreasing in n.

It makes intuitive sense that the easier one can copy the leaders even from a distance, the less incentive one has to keep up with them.[3]

V. Empirical Implications

We end the paper by describing and testing some implications and estimating $\alpha(\cdot)$.

A. *Firm Size and TFP*

A central premise of (11) is that $A'(k) < 0$. That is, the leaders have lower total factor productivity (TFP) than the followers. In equilibrium, the elasticity of TFP with respect to k is $-\varepsilon(k)$. Equation (5) and the constancy of $\varepsilon(k)$ (in equilibrium) with respect to k means that in equilibrium $A(k) = A_0 k^{-\varepsilon}$, and that the production function looks like

$$y = A(k)k = A_0 k^{1-\varepsilon}.$$

This means that if we were to estimate returns to scale, or the "span of control," we should find that the elasticity of output is $1 - \varepsilon$.

On the face of it, this implication agrees with the bulk of the empirical findings on TFP and firm size. First, large firms are less productive in research than small firms. Per R&D dollar spent, research output declines with firm size (see Wesley M. Cohen and Steven Klepper, 1992).

Second, small plants seem to be more productive than large plants. G. Steven Olley and Ariel Pakes (1996) find decreasing returns and estimate elasticity of output with respect to inputs is around 0.95, which would imply an ε of around 0.05. Andrew Atkeson and Patrick J. Kehoe (2001) find that ε is around 0.04. Some estimates of ε are larger than that (e.g., Atkeson et al., 1996), while others are smaller (e.g., Susanto Basu and John G. Fernald [1997] estimate ε at around 0.01–0.02). Our own estimates of ε, obtained in a variety of ways, will lie between 0.013 and 0.056.[4]

B. *Estimation with Stock-Market Data*

Our model explains the distribution of firm values with a spillover mechanism. In fact, one may be able to identify the spillover mechanism from the observed distribution of firm size. We now show this in the context of firms' stock-market values. Our propositions are about a steady state, and so we will assume that all firms, and hence the distribution, are on the balanced growth path, with stationary growth factor x.

Since in equilibrium we must have $A_t(k) = \lambda_t k^{-\varepsilon}$ and since $\lambda_t = K_t^{\beta+\varepsilon}$, the production function, as it would appear in a sample of firms that are in a long-run growth equilibrium, is

$$y = k^{1-\varepsilon} K_t^{\beta+\varepsilon}. \tag{25}$$

In steady state, the firm's dividend is $[1 - C(x)]y$, and so, since all capital grows by a factor x, the present value of its dividends at date t is

$$\begin{aligned} v_t(k) &= (1 - C)\sum_{\tau=t}^{\infty} \frac{(x^{\tau-t}k)^{1-\varepsilon}(x^{\tau-t}K_t)^{\beta+\varepsilon}}{(1+r)^{\tau-t}} \\ &= (1 - C)\left[1 - \frac{x^{\beta}}{1+r}\right]^{-1} K_t^{\beta+\varepsilon} k^{1-\varepsilon} \end{aligned} \tag{26}$$

[3] We thank a referee for suggesting this formulation of the access function.

[4] Unfortunately, the model cannot, it seems, explain inequality among countries. Robert E. Hall and Charles I. Jones (1999, Table 1) find that, among countries, the level of total factor productivity varies *positively* with income. Some models of countries that focus on diffusion lags are Stephen L. Parente and Edward C. Prescott (1994), Robert J. Barro and Xavier Sala-i-Martin (1997), and Basu and David N. Weil (1998).

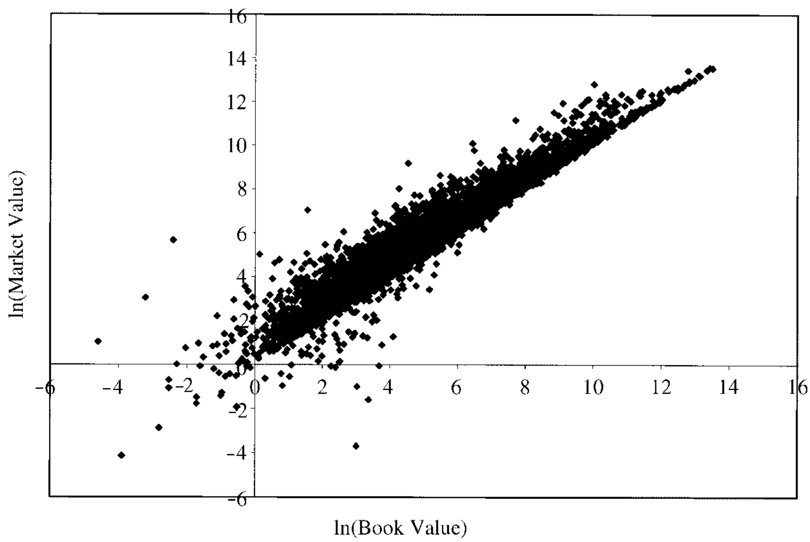

FIGURE 4. THE SCATTERPLOT OF BOOK VALUE AGAINST MARKET VALUE (ln SCALE)

since in steady state, $K_{t+1} = (1 + x)K_t$. Therefore we arrive at the equation

$$\text{(27)} \qquad \ln v = \pi_t + (1 - \varepsilon)\ln k$$

where

$$\pi_t = \ln\left\{(1 - C)\left[1 - \frac{x^\beta}{1 + r}\right]^{-1} K_t^{\beta + \varepsilon}\right\}$$

is a constant that does not depend on k. Note that in general, $C(k_{t+1}/k_t)$ depends on k. However, in the steady state $k_{t+1}/k_t = x$ for all k.

Identifying ε.—To estimate equation (27), we use Compustat data, a cross-section sample of 8,276 firms in 1998. For a firm's k we use book values of capital; that is, historical values of the stock of purchases of capital using some measure of depreciation. For v we use the firm's market value. We assume that these firms were in a long-run equilibrium at the moment in time (i.e., end-of-year, 1998) that we measure these variables, and we use only cross-sectional information. The estimates are

$$\text{(28)} \qquad \ln v = \underset{(0.018)}{1.099} + \underset{(0.003)}{0.944} \ln k$$

($R^2 = 0.901$, $n = 8{,}276$), which gives a point estimate $\hat{\varepsilon} = 0.056$ significantly different from zero. Figure 4 gives the scatterplot of $\ln(v)$ on $\ln(k)$. This is consistent with the presence of free-riding based on spillovers.

Identifying β.—This parameter is identifiable in the usual way—from the growth of TFP as Paul M. Romer (1986) and Lucas (1988) have argued. If a firm's k is on its books, and if we do the usual Solow accounting that assumes constant returns to scale at the firm level, a firm's TFP growth is, (using g_i to denote the growth rate of variable i),

$$g_y - g_k = (1 - \varepsilon)g_k + (\beta + \varepsilon)g_K = \beta g_K = \beta(x - 1)$$

because k grows at the same rate as K. Assuming an aggregate production function that includes physical and human capital as components of k, Lucas (1988) has used this growth-accounting method to conclude that $\hat{\beta} = 0.4$. We shall use his estimate.

Identifying $\alpha(z)$.—We can use (19), which tells us that, in equilibrium,

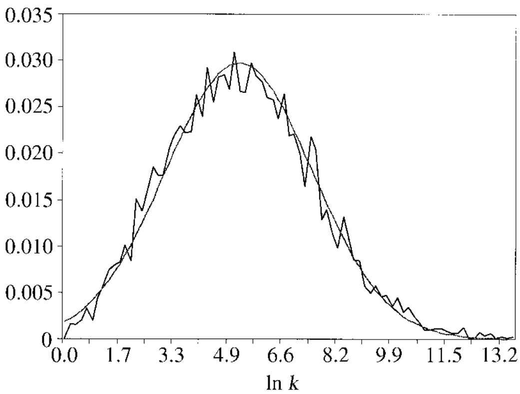

FIGURE 5. THEORETICAL (NORMAL) AND EMPIRICAL PROBABILITY DENSITY FUNCTION OF ln k

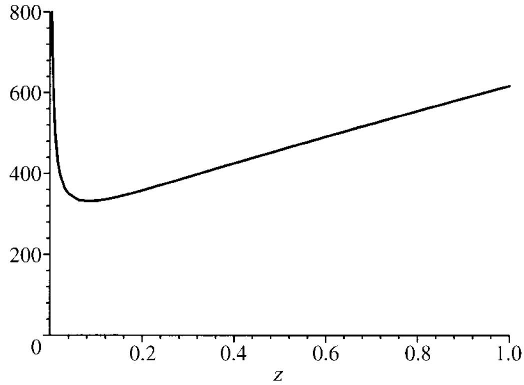

FIGURE 6. THE ESTIMATED COPYING TECHNOLOGY FUNCTION $\hat{\alpha}$

$$\text{(29)} \qquad \alpha(z) = \frac{\varepsilon z^{-1-\varepsilon/\beta}}{\beta h(z)}$$

and (26), which tells us that market value is proportional to k. The empirical distribution of book values k appears to be lognormal, and we shall fit to it a lognormal density parameterized as

$$\phi(k) = \frac{1}{\sqrt{2\pi}k\sigma} e^{-(\ln k - \mu)^2/2\sigma^2}.$$

The maximum-likelihood estimators for this mean and the variance of this lognormal distribution are $\hat{\mu} = 5.25$ and $\hat{\sigma} = 2.23$. The empirical and theoretical probability density functions are plotted in Figure 5.

A Kolmogorov-Smirnov test confirms the hypothesis that the fitted distribution, $\hat{\phi}(k)$, coincides with the empirical distribution. The Kolmogorov-Smirnov test statistic is 0.0121 and significant (at the 5-percent level, the critical value is 0.0166), so both distributions are statistically identical.

Now, given k is lognormally distributed, we can apply the change of variables theorem to $\phi(k)$ to obtain $h(z)$ (where $z = k/K$). This is the transformation of the lognormal distribution (with domain $\mathbb{R}_+$) into the corresponding distribution with domain $[z_m, 1]$:

$$h(z) = \frac{1}{\sqrt{2\pi}z\sigma} e^{-(\ln K - \mu + \ln z)^2/2\sigma^2}.$$

Now using equation (29), we substitute $h(z)$ to obtain $\alpha(z)$:

$$\alpha(z) = \frac{\varepsilon}{\beta} z^{-\varepsilon/\beta} \sqrt{2\pi}\sigma e^{(\ln K - \mu + \ln z)^2/2\sigma^2}.$$

Any copying technology of the form $\alpha(z) = c_1 z^{c_2} \exp\{(c_3 + \ln z)^2/c_4\}$ (where the c_i are constants) generates an equilibrium distribution of k that is lognormal. We can use the estimates for $\hat{\varepsilon}$, $\hat{\beta}$, $\hat{\sigma}$, $\hat{\mu}$, to get $\hat{\alpha}$. From the data, we also find that the log of the largest firm's capitalization, $\ln K = 13.4$.

What can we learn about the copying technology? Given the result of Proposition 6, and given that firms' stock-market valuations are highly dispersed, we would not expect the search process to be very directed. And while high-z firms can, to an extent, direct their copying toward the leaders, copying by low-z firms seems to be directed *away* from the leaders. Figure 6 plots $\hat{\alpha}$ which is U-shaped. This indicates that a lot is copied in the immediate neighborhood of the leaders, and also that firms in the middle region find it harder to copy the leaders. At the lower end of the distribution, firms find it much easier to copy their immediate neighbors than to copy the leaders.

Is this conclusion reasonable? We can draw a parallel between how producers copy other producers on the one hand, and how inventors cite other inventors on the other. When an academic writes a paper, he cites work on which he has built or work that he wishes to distinguish from his own, even though it is similar to his. Similar motives guide an inventor when she applies for a patent. She cites related patents, and patents that she wishes to distinguish from her own. Occasionally, the patent examiner adds citations of his own. In this context, citations are directed toward the frontier if the most recent and, presumably, state-of-the-art patents are cited.

C. *Evidence on Patent Citation*

In our model, as in Lucas (1988), copying enhances a firm's productivity, A, but does not raise the firm's capital stock. Suppose that

1. Each instance of copying is accompanied by a patent citation. If you use someone's idea, you must cite it.
2. A patent's "quality" is the z of the firm that owns it, so that the distribution of patent quality over firms is $h(z)$.
3. The probability that a patent z is cited is proportional to $\alpha(z)$.
4. A patent's z falls with its age, τ, as given by

$$z = x^{-\tau}$$

where $x = k'/k \geq 1$ is the growth factor common to all firms.

Then equation (11) suggests that firm s will cite the patents of the firms that are more productive than it is, and that

$$\Pr(s \text{ cites } z) = \begin{cases} \alpha(z)h(z) & \text{if } s < z \\ 0 & \text{if } s \geq z. \end{cases}$$

This says, as before, that firm s, will run into a firm of type z with probability $h(z)$, and the probability that firm s will want to, or will be able to copy that firm is $\alpha(z)$. As before, we assume that firm s wants to copy only firms that are better than itself. Therefore the expected number of citations that a firm z will get is

$$N(z) = \alpha(z) \int_{z_m}^{z} h(s)\, ds$$

$$= \alpha(z)H(z)$$

for $z \in [z_m, 1]$. Because imitation is not perfect, $\alpha(z)$ is positive even for values of z significantly below 1, and hence, each invention is copied for a long time after it has been patented. Let τ be the time elapsed since an invention was patented. The frontier technology has been patented in the current period: for $z = 1$, we have $\tau = 0$. Inside the frontier, the lower the technology z, the older is the patent (i.e., the higher is τ). Now using the change of variables theorem, we can derive the number of citations $\bar{N}(\tau)$ as a function of age. Assumption 4 implies that $dz/d\tau = x^{-\tau} \ln x$:

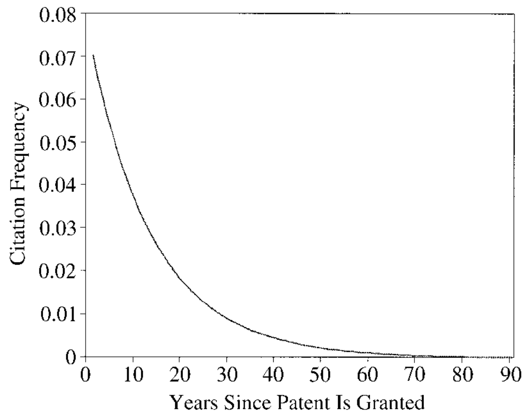

FIGURE 7. THE PREDICTED NUMBER OF CITATIONS AS A FUNCTION OF PATENT AGE τ: $\bar{N}(\tau)$

$$\bar{N}(\tau) = N(z)\,\frac{dz}{d\tau}$$

$$= x^{-\tau}\ln(x)N(x^{-\tau})$$

$$= x^{-\tau}\ln(x)\alpha(x^{-\tau})H(x^{-\tau}).$$

Then, using the estimated $\hat{\alpha}(z)$,

$$\hat{\alpha}(z) = \frac{\hat{\varepsilon}}{\hat{\beta}}\, z^{-\hat{\varepsilon}/\hat{\beta}} \sqrt{2\pi}\hat{\sigma} e^{(\ln \hat{K} - \hat{\mu} + \ln z)^2/2\hat{\sigma}^2}$$

where $\ln \hat{K} = 13.4$, $\hat{\varepsilon} = 0.056$, $\hat{\beta} = 0.4$, $\hat{\sigma} = 2.23$, and $\hat{\mu} = 5.25$, and using the estimated empirical distribution $H(z)$ (lognormal) and a growth factor $x = 1.03$, we can now plot the predicted number of citations as a function of the age of the patent: $\bar{N}(\tau)$ (see Figure 7).

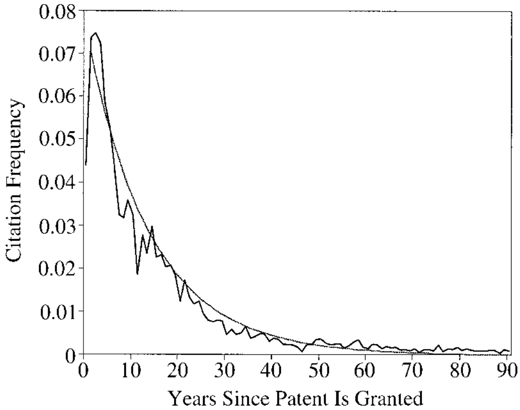

FIGURE 8. THEORETICAL AND EMPIRICAL AGE DISTRIBUTION OF CITATIONS

We use data from Caballero and Jaffe (1993) on citations for the year 1991. Their data consist of a 1 in 100 random sample of all patents in the United States granted in that year. Their sample reports data on a total of 6,961 patents. In addition, they have observations on the number of times older patents have been cited in their 1991 sample of patents. For example, all of their 1991 sample patents cite 364 patents granted in 1985, they cite 32 patents granted in 1960, and they cite 8 patents granted in 1901. In Figure 8, we have plotted the empirical frequency of citations in 1991 against the number of years since the cited patent was granted (starting in 1990 up to 1901; i.e., during 90 years). Note that the mode of the empirical distribution in 1991 is equal to 2.[5] In addition, there is a plot of the theoretical distribution of citations.[6]

We perform a chi-square test which yields a test statistic of 0.539, so we cannot reject the hypothesis that the theoretical and empirical distributions are identical! A remarkable coincidence given that $\hat{\alpha}$ was estimated using data other than citations.

D. *Discussion: Possible Bias in* $\hat{\varepsilon}$

The specification in (28) may lead to a biased estimate of ε for two reasons. First, the model may leave out some variables that influence v and that may be correlated with our measure of k and thereby lead to a biased estimate in its coefficient. And second, capital, k, may itself be poorly measured, leading to a classic errors-in-variables bias toward zero in the coefficient of k.

1. ***Bias Due to Unmeasured Capital.***—Ellen R. McGrattan and Prescott (2000) argue that as much as 40 percent of GNP is intangible capital. An unmeasured component of k would affect v and, if it is correlated with the measured part of k, our estimate of ε would be biased. We shall use two separate proxies of intangibles.

Age as a Proxy for Intangible Capital.—Byong-Hong Bahk and Michael Gort (1993) find that plant TFP rises with the age of the plant. Old plants and, therefore, old firms are more productive because they have accumulated the intangible expertise. But while age should raise the firm's TFP, it lowers the future growth of that TFP, and this is what matters for the interpretation of the estimates in (28). Roughly speaking, a young firm has less capital on its books, and a sharply rising TFP trajectory compared to other firms. The firm's market value capitalizes the growth in TFP and the young firm will have a higher market/book ratio. If this argument is correct, the estimated coefficient of ln k in (28) would then be less than unity, but not for the reasons implied by our model. To eliminate this potential bias, we use two different subsamples of the Compustat data, selected on the availability of observations on age.[7] The INC sample has observations on the age since incorporation of the firm. The LIST sample has observations on the age since the firm's stock exchange listing. In Table 1 we report the regressions. Column (1) is the regression reported earlier on the original sample. Columns (2) and (4) repeat the same regression

[5] For the entire data set, Caballero and Jaffe (1993, p. 34) report that: "... the distributions over [τ] have extremely long tails. The mean lag in years is about 16 years; the median is about 10, and the mode is about 3."

[6] Both are normalized to be probability distributions with measure one.

[7] The age sample was analyzed by Jovanovic and Rousseau (2001b) who provide a full description of it.

TABLE 1—USING AGE TO PROXY FOR INTANGIBLES

	Original Sample	INC Sample		LIST Sample	
Variable	(1)	(2)	(3)	(4)	(5)
ε	0.056	0.014	0.013	0.034	0.037
	(0.003)	(0.006)	(0.006)	(0.004)	(0.004)
ln[age]	—	—	−0.008[a]	—	−0.030
			(0.011)		(0.007)
Intercept	1.099	0.918	0.938	0.979	1.020
	(0.018)	(0.035)	(0.044)	(0.020)	(0.007)
R^2:	0.901	0.910	0.910	0.913	0.913
Number of observations:	8,276	2,898		6,265	

Note: Standard errors are in parentheses.
[a] Not significant at 10-percent significance level.

for each of the subsamples. Regressions (3) and (5) include ln(age) as a regressor.

The data-reporting requirement may have excluded, especially from the INC sample, some firms that were found in "high ε" activities, for want of a better phrase. Before including age, the estimates of ε are lower and significant, and highly so in the LIST sample. Interestingly, in both samples, the inclusion of the age variable diminishes neither the estimate of ε nor its level of significance. Although age is significant only in the LIST sample, the coefficient is, in both samples, negative, as we expected. Moreover, the estimate of ε is still very much in the range of the studies described above.

Reported Intangibles as a Regressor.—The Compustat data contain a variable called "intangibles." It includes blueprints, client lists, patent costs and copyrights, goodwill, trademarks and tradenames, but excludes software. The regression of market value on book and intangibles is reported here for a sample of firms for which the "intangibles" variable is available:

$$\ln v = \underset{(0.028)}{1.107} + \underset{(0.008)}{0.922} \ln k + \underset{(0.007)}{0.038} \ln(\text{intangibles})$$

($R^2 = 0.916$, $n = 4{,}270$). The intangibles variable is highly significant, but the estimate of ε,

$$\hat{\varepsilon} = 1 - 0.922 - 0.038 = 0.040$$

is roughly the same as before, and still significantly different from zero. Without intangibles in the regression, this sample yields almost the same estimate $\hat{\varepsilon} = 0.041$

$$\ln v = \underset{(0.025)}{1.026} + \underset{(0.004)}{0.959} \ln k$$

($R^2 = 0.915$, $n = 4{,}270$).

A Possible Downward Bias on $\hat{\varepsilon}$.—So far, we discussed possible upward bias in $\hat{\varepsilon}$ (i.e., downward bias in the coefficient of ln k). But note that there are reasons to expect that $\hat{\varepsilon}$ is, in fact, biased *downward.* That is, the coefficient of k may be biased upward because of reverse causality. A shock to the firm's production function would raise v, and this is the numerator in Tobin's Q. The Q theory says that k will rise in response, and thus one would overestimate the coefficient on k in this context.

2. ***An Alternative Estimate of*** ε.—If k is poorly measured, we can use the Compustat data on sales to estimate the model. Observe that in a cross section of firms, (27) implies that

$$\ln\left(\frac{v}{k}\right) = \pi_t - \varepsilon \ln k \tag{30}$$

while (25) implies that

$$k = (K^{-\beta-\varepsilon} y)^{1/(1-\varepsilon)}. \tag{31}$$

Then (30) and (31) imply that

$$\ln\left(\frac{v}{k}\right) = \bar{\pi}_t - \frac{\varepsilon}{1-\varepsilon}\ln y$$

where

$$\bar{\pi}_t = \pi_t - \frac{\varepsilon(\beta+\varepsilon)}{1-\varepsilon}\ln K_t.$$

Once again we use the 1998 cross-section Compustat sample, using market-to-book ratio for v/k and sales for y. The regression comes out as

$$\ln\left(\frac{v}{k}\right) = \underset{(0.017)}{0.977} - \underset{(0.003)}{0.030}\ln y$$

($R^2 = 0.0099$, $n = 8{,}276$). The coefficient on $\ln y$ is negative,[8] highly significant and implies the point estimate

$$\hat{\varepsilon} = \frac{0.03}{1.03} = 0.029$$

which is lower than the original estimate of $\hat{\varepsilon} = 0.056$, but still significantly positive.

VI. Conclusion

Spillovers have, for a long time, been associated with free-riding and with an equilibrium effort below the social optimum. And, since spillovers flow from leaders to followers, they have, so far, been held to be a force of convergence toward equality. In this paper, we show that the prospect of receiving spillovers may induce followers to relax their efforts so much that, instead of diminishing, inequality will rise.

We solved for the equilibrium size distribution in terms of the spillover process at the micro level. The larger the spillovers and the larger the benefits of hanging back, the greater the inequality. Inequality rises even further if copying can be directed toward the leaders. We showed that these properties hold generally, and we provided examples.

Finally, we estimated the spillover mechanism from stock-market data and found that it was highly undirected. We confirmed this finding in patent-citations data where undirected copying shows up as citation of old patents. The overall impression is that firms cannot copy one another nearly as easily as we sometime assume to be the case. If they could, our model suggests that the dispersion of firm sizes would be even greater than it now is.

APPENDIX

Defining $A_t(k)$ for $k > K$.—The definition of $A_t(k)$ over the entire support satisfies

$$A_t(k) = \begin{cases} \left[K_t\left(1 + \int_{k/K}^{1} \alpha(z)h_t(z)\,dz\right)\right]^{\beta} & \text{if } k \le K \\ \left[K_t\left(1 + \alpha(1)h(1)\left[1 - \frac{k}{K_t}\right]\right)\right]^{\beta} & \text{if } k > K. \end{cases}$$

Then for the first derivative of A,

$$A'(k) = \begin{cases} -\beta A_t(k)^{\beta-1}\alpha\left(\frac{k}{K}\right)h_t\left(\frac{k}{K}\right) & \text{if } k \le K \\ -\beta A_t(k)^{\beta-1}\alpha(1)h(1) & \text{if } k > K. \end{cases}$$

Note that under this specification, $A'(k)$ is continuous at $k = K$, and that the production set is convex for any feasible $k \in \mathbb{R}_+$.

*Example 2: $\alpha(z) = \alpha/z$ (*copying directed away from the leaders*).*

Equation (19) implies that

$$h(z) = \frac{\varepsilon}{\alpha\beta}z^{-\varepsilon/\beta} \qquad \text{for } z \in [z_m, 1]$$

where, from (20) it follows that

$$z_m = \left[1 + \alpha - \frac{\alpha\beta}{\varepsilon}\right]^{\beta/(\beta-\varepsilon)}.$$

This solution is valid for $0 < \alpha\beta < \varepsilon(1+\alpha)$. This solution for z_m has the same features as example 1. As α or β go to zero, z_m tends to unity, and inequality disappears.

[8] The finance literature (e.g., Karl Lins and Henri Servaes, 1999) refers to this relation as the "size discount."

PROOF OF LEMMA 1:

The first derivative of $\Psi(x)$ with respect to x is

$$\frac{d}{dx}\left\{\left[\frac{1}{\rho}x^{(1+\beta)(\gamma-1)+1}-1\right]\frac{C'(x)}{1-C(x)}\right\}$$

$$=\left[\frac{1}{\rho}((1+\beta)(\gamma-1)+1)x^{(1+\beta)(\gamma-1)}-1\right]$$

$$\times\frac{C'}{1-C}+\left[\frac{1}{\rho}x^{(1+\beta)(\gamma-1)}-1\right]$$

$$\times x\,\frac{C''(1-C)+(C')^2}{[1-C]^2}.$$

For an equilibrium to exist, consumption must not grow faster than the interest rate: $1+r>x_c$, which implies

$$1+r>x_kx_A=x^{1+\beta}.$$

Then it follows that

$$\frac{1}{\rho}x^{(1+\beta)(\gamma-1)}>1.$$

That implies that the second term is always positive (note that $1-C>0$). Now as a result, for any $\gamma\geq 1$, the first term is positive as well.

Note that Lemma 2 holds for a more general class (i.e., $\gamma\geq 1$ is a sufficient, but not a necessary, condition). A more general characterization and proof can be obtained from the authors upon request.

REFERENCES

Atkeson, Andrew and Kehoe, Patrick J. "Measuring Organization Capital." Staff Report 291, Federal Reserve Bank of Minneapolis, July 2001.

Atkeson, Andrew; Khan, Aubhik and Ohanian, Lee. "Are Data on Industry Evolution and Gross Job Turnover Relevant for Macroeconomics?" *Carnegie-Rochester Conference Series on Public Policy*, June 1996, *44*, pp. 216–50.

Bagnoli, Mark and Lipman, Barton. "Provision of Public Goods: Fully Implementing the Core through Private Contributions." *Review of Economic Studies*, October 1989, *56*(4), pp. 583–601.

Bahk, Byong-Hong and Gort, Michael. "Decomposing Learning by Doing in New Plants." *Journal of Political Economy*, August 1993, *101*(4), pp. 561–83.

Barro, Robert J. and Sala-i-Martin, Xavier. "Technological Diffusion, Convergence, and Growth." *Journal of Economic Growth*, March 1997, *2*(1), pp. 1–26.

Basu, Susanto and Fernald, John G. "Returns to Scale in U.S. Production: Estimates and Implications." *Journal of Political Economy*, April 1997, *105*(2), pp. 249–83.

Basu, Susanto and Weil, David N. "Appropriate Technology and Growth." *Quarterly Journal of Economics*, November 1998, *113*(4), pp. 1025–54.

Bénabou, Roland. "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance." *American Economic Review*, June 1996, *86*(3), pp. 584–609.

Caballero, Ricardo and Jaffe, Adam. "How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," in Oliver Jean Blanchard and Stanley Fischer, eds., *NBER macroeconomic annual 1993*. Cambridge, MA: MIT Press, 1993, pp. 15–74.

Cohen, Wesley M. and Klepper, Steven. "The Anatomy of Industry R&D Intensity Distributions." *American Economic Review*, September 1992, *82*(4), pp. 773–99.

Gabaix, Xavier. "Zipf's Law for Cities: An Explanation." *Quarterly Journal of Economics*, August 1999, *114*(3), pp. 739–67.

Galor, Oded and Tsiddon, Daniel. "Technological Progress, Mobility, and Economic Growth." *American Economic Review*, June 1997, *87*(3), pp. 363–82.

Glomm, Gerhard and Ravikumar, B. "Public versus Private Investment in Human Capital Endogenous Growth and Income Inequality." *Journal of Political Economy*, August 1992, *100*(4), pp. 813–34.

Hall, Robert E. and Jones, Charles I. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics*, February 1999, *114*(1), pp. 83–116.

Ioannides, Yannis M. and Overman, Henry G. "Zipf's Law for Cities: An Empirical Examination." Mimeo, Tufts University, 2000.

Jovanovic, Boyan and Lach, Saul. "Entry, Exit, and Diffusion with Learning by Doing." *American Economic Review*, September 1989, *79*(4), pp. 690–99.

Jovanovic, Boyan and MacDonald, Glenn M. "Competitive Diffusion." *Journal of Political Economy*, February 1994, *102*(1), pp. 24–52.

Jovanovic, Boyan and Rousseau, Peter L. "Vintage Organization Capital: 1885–1998." Mimeo, University of Chicago, 2001a.

______. "Why Wait? A Century of Life Before IPO." *American Economic Review*, May 2001b (*Papers and Proceedings*), *91*(2), pp. 336–41.

Lins, Karl and Servaes, Henri. "International Evidence on the Value of Corporate Diversification." *Journal of Finance*, December 1999, *54*(6), pp. 2215–39.

Lucas, Robert E., Jr. "On the Mechanics of Economic Development." *Journal of Monetary Economics*, July 1988, *22*(1), pp. 3–42.

McGrattan, Ellen R. and Prescott, Edward C. "Is the Stock Market Overvalued?" *Federal Reserve Bank of Minneapolis Quarterly Review*, Fall 2000, *24*(4), pp. 20–40.

Nelson, Richard R. "Modelling the Connections in the Cross Section between Technical Progress and R&D Intensity." *RAND Journal of Economics*, Autumn 1988, *19*(3), pp. 478–85.

Olley, G. Steven and Pakes, Ariel. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, November 1996, *64*(6), pp. 1263–97.

Parente, Stephen L. and Prescott, Edward C. "Barriers to Technology Adoption and Development." *Journal of Political Economy*, April 1994, *102*(2), pp. 298–321.

Romer, Paul M. "Increasing Returns and Long-run Growth." *Journal of Political Economy*, October 1986, *94*(5), pp. 1002–37.

Stigler, George J. "The Economics of Information." *Journal of Political Economy*, June 1961, *69*(3), pp. 213–25.

Tamura, Robert. "Income Convergence in an Endogenous Growth Model." *Journal of Political Economy*, June 1991, *99*(3), pp. 522–40.

Tiebout, Charles. "A Pure Theory of Local Expenditures." *Journal of Political Economy*, October 1956, *64*(5), pp. 416–24.