GROWTH THROUGH HETEROGENEOUS INNOVATIONS

Ufuk Akcigit
William R. Kerr

Growth Through Heterogeneous Innovations
Ufuk Akcigit and William R. Kerr
NBER Working Paper No. 16443
November 2010
JEL No. L16,O31,O33,O41

<u>ABSTRACT</u>

We study how exploration versus exploitation innovations impact economic growth through a tractable endogenous growth framework that contains multiple innovation sizes, multi-product firms, and entry/exit. Firms invest in exploration R&D to acquire new product lines and exploitation R&D to improve their existing product lines. We model and show empirically that exploration R&D does not scale as strongly with firm size as exploitation R&D. The resulting framework conforms to many regularities regarding innovation and growth differences across the firm size distribution. We also incorporate patent citations into our theoretical framework. The framework generates a simple test using patent citations that indicates that entrants and small firms have relatively higher growth spillover effects.

Ufuk Akcigit
Department of Economics
University of Pennsylvania
3718 Locust Walk, #445
Philadelphia, PA 19104
and NBER
uakcigit@econ.upenn.edu

William R. Kerr
Harvard Business School
Rock Center 212
Soldiers Field
Boston, MA 02163
and NBER
wkerr@hbs.edu

1 Introduction

Despite the rapid development of endogenous growth theory over the last two decades, we still have very little understanding about the innovation and growth patterns of different-sized firms and their contributions to macroeconomic growth. The motivations of small start-ups for R&D and innovation differ greatly from Fortune 100 conglomerates, however, and our understanding of economic growth and appropriate innovation policies will be limited until we shed light on these key differences. An important step forward is to characterize the theoretical and empirical relationships between firm size and R&D investment behavior, which then determines innovation and growth dynamics.

In this study, we make a distinction between two types of R&D that firms undertake: *exploration* and *exploitation*. Firms undertake exploration R&D to create new products and capture leadership in markets. On the other hand, firms undertake exploitation R&D to improve product lines that they are currently serving. We begin by showing some basic empirical regularities regarding this important distinction, which has been under explored so far. Building on the seminal work of Klette and Kortum (2004), we then introduce this key distinction into a tractable endogenous growth framework and generate many important, testable predictions regarding firm size and R&D spending, patenting, patent citations, and growth dynamics. Our final task is to test these predictions using Census Bureau data.

Our paper makes several key theoretical advances. Most importantly, we offer a micro-foundation for two types of R&D in a general equilibrium framework with firm entry and exit. In our model, firms invest in R&D for two reasons: to improve their mark-ups on their existing products and to innovate new products. The model provides a rich characterization of the firm dynamics due to heterogenous R&D and innovation behavior across different-sized firms. In other words, it allows small start-ups to differ in meaningful ways from General Electric and Procter & Gamble. We are not aware of prior work that has built this degree of heterogeneity in R&D choices into a tractable framework with closed-form solutions.

Second, incorporating this exploration-exploitation distinction allows us to connect to several empirical moments that vary across the firm size distribution. Examples include deviations from Gibrat's Law (i.e., small firms grow faster conditional on survival), declines in exploration R&D shares with firm size, and declines in conditional R&D intensities with firm size among innovative firms. Prior models typically generate moments like growth rates and R&D intensities that are uniform across the distribution. Our model allows for substantially greater heterogeneity among firms and generates interesting distributional implications.

Our third contribution is to explicitly incorporate patent citations into the endogenous growth framework. These additions allow us to assess the link between firm size and its citations given to and received from other firms. We derive tests that employ patent citations to determine if the growth spillover effects from exploration innovations are larger or smaller than those from exploitation innovations, which is a crucial input to optimal policy design. Finally, our model conceptualizes technology clusters. These clusters provide an intuitive foundation for declining citation rates with invention age, which are not generated in basic endogenous growth models.

Section 2 formalizes our theoretical model. We first consider a benchmark case where exploration R&D does not scale with firm size at all. Under these conditions, we derive very tractable analytical solutions of the dynamics of economic growth and the concomitant heterogeneity among innovative firms. We then discuss how our results hold in an alternative

setting with partial scaling of exploration R&D. Using an exploration R&D technology that provides a size advantage to large firms a là Klette and Kortum (2004), we further show that our benchmark results hold also with this alternative technology if firms also face managerial costs with larger firm size. At the end of Section 2, we list the predictions of our model regarding firm entry/exit, R&D, growth, innovation, patent counts, and patent citations.

Section 3 tests these empirical predictions. We show that our model matches many known empirical facts regarding the firm size distribution such as the high skewness of the distribution. Moreover, our model has sharp predictions on some long-standing debates: It predicts that small firms grow faster and that their R&D to sales ratio is higher than large firms on average. These two predictions are strongly supported by our data for the recent US economy. We further show empirical evidence regarding the model's unique predictions. For instance, we show that the relative rate of major inventions is higher in small firms and that the exploitation innovation share is higher in large firms. We demonstrate that these distributional differences are not due to differences in research capabilities or technologies, but are instead an outcome of R&D investment choices by firms.

Implementing our derived patent citations test, we find that growth spillover effects are larger from exploration R&D than exploitation R&D, akin to first-order stochastic dominance. Our model identifies a comparative advantage for new entrepreneurs and small firms in undertaking exploration innovation due to increased managerial attention of large firms on refining their existing product lines. Under the conditions identified by our test, small firms and new entrants also yield greater spillovers, in a relative sense, into economic growth due to these forces.

Section 4 concludes. Our work connects two literatures that have largely remained separate. First, we clearly build on a lengthy theoretical growth literature.[1] Our framework offers an intuitive, tractable way of bringing more of the realities of the empirical literature into the workhorse theoretical models in the spirit of Klette and Kortum (2004) and Akcigit (2010). Second, a large body of work examines the empirics of innovation. Section 3 outlines the antecedents of our study from this literature, with particular emphasis on the heterogeneity of innovation size and its relationship to the firm size distribution (e.g., Acs and Audretsch 1987, 1988, Kortum and Lerner 2000, Baumol 2009).

It is important to note that we focus solely on innovative firms in this paper, which we define to be firms undertaking R&D activity and patenting. This concentration allows us to connect a closed-form endogenous growth model that embraces firm heterogeneity with parallel empirical work. The firm dynamics of this sector are of direct policy interest (e.g., support to small R&D firms), and innovation is a central driver for overall economic growth. In parallel efforts, we are numerically quantifying a version of the model that includes non-innovative firms at the expense of losing analytical tractability. We likewise hope that future researchers find our framework to be a useful platform for studying firm heterogeneity and economic growth.

---

[1] For example, Romer (1986, 1990), Grossman and Helpman (1991), Aghion and Howitt (1992), Aghion et al. (1997, 2001), Barro and Sala-i-Martin (1995), Kortum (1997), Klette and Kortum (2004), and Hopenhayn et al. (2006). Related papers on heterogeneous innovation are Akcigit (2010), Acemoglu and Akcigit (2010), Acemoglu and Cao (2010), Cai (2010), and Acemoglu et al. (2010). Related papers on spillover benefits are Spence (1984) and Griliches (1992). Caballero and Jaffe (1993) also connect patent citations with growth models. Acemoglu (2009) provides a comprehensive overview.

Empirical Foundations

Before proceeding to the model, it is helpful to introduce several empirical regularities and their theoretical counterparts. Our exploration-exploitation terminology is extensively used in the management and organizational behavior literatures, and it has clear parallels with the pivotal work of Klepper (1996) and Cohen and Klepper (1996b) on product versus process R&D. This latter work demonstrates that the share of R&D directed towards process improvements grows with firm size and industry maturity. They model how this pattern descends from process innovations being less saleable in disembodied form compared to product innovations. This results in the returns to process R&D growing with firm size.

Figure 1 documents Census Bureau data for innovative firms. Each panel describes an aspect of exploitation R&D. Panels 1A and 1B consider the extent to which patents of each firm cite the firm's prior work. High self citation rates are indicative of exploitation R&D that is meant to enhance the firm's current technologies. Panels 1C and 1D similarly consider the extent to which the firm engages in process-oriented R&D, a proxy for exploitation R&D.

The three groups of columns per panel separate firms by employment size. Within each triplet, firms are further separated by contemporaneous employment growth. We use three buckets: firms that are exiting or showing major employment declines (gray bars), continuing firms that show small or modest employment changes (black bars), and firms that are entering or showing major employment gains (white bars). Major expansions or declines in employment are demarcated by a 33% change over two years around the survey year.[2]

The first regularity is that exploitation behavior grows in firm size. This is particularly true on the extensive margin of Panels 1A and 1C, and it also holds on the intensive margin. These share growths are very regular and hold within more finely demarcated size groups. Large firms engage in relatively more exploitation R&D than small firms, such that exploration R&D shares are inversely related to firm size. Our model generates these features as firms choose R&D strategies optimal for their portfolio of product lines. Firms with more incumbent operations devote relatively more attention to improving these operations.

The second regularity is that exploitation R&D, by contrast, is approximately invariant to contemporaneous firm growth rates. This relationship is seen by comparing across the columns within each triplet. There is much more variation across the firm groups than within the triplets. Exploitation R&D depends substantially on firm size but cannot be predicted by contemporaneous growth rates once firm size is controlled for. In our model, realized outcomes of innovative efforts are stochastic conditional on the type of R&D undertaken. This randomness in outcomes yields contemporaneous growth rates that are not further related to the type of innovation pursued except as determined by firm size.

In Panel 1B, we list above the columns the shares of innovative firms in our sample. Small firms are greater in number than large firms and show higher variability in growth rates. These are well known attributes of the firm size distribution, and our model generates these features due to the relatively larger impact of adding or losing product lines for small firms.

---

[2]Innovation data come from the NSF Survey of Industrial Research and Development and the NBER Patent Database. We describe our data and its preparation in Section 3 and the appendix. Section 3 also provides Monte Carlo simulations of self citation behavior and firm size. Figure 1 restricts the sample to years in which product versus process R&D are reported (i.e., odd years from 1979 to 1989).

For a year t, we define firm employment levels as the average of non-zero employment from $t-1$ to $t+1$. We define employment growth as $[Emp_{t+1} - Emp_{t-1}]/[(Emp_{t+1} + Emp_{t-1})/2]$. This growth measure is bounded by (-2,2) for continuing firms. Entrants and exiting firms have a value of 2 and -2, respectively.

2 Theoretical Framework

We now introduce our model that incorporates the empirical regularity that exploration R&D does not scale as fast as exploitation R&D with firm size. Our goal is to study the implications of this heterogeneity on the R&D, innovation, and growth dynamics of firms. We first describe the model's basic environment and then solve for its steady state equilibrium in a setting where exploration R&D does not scale with firm size. We then show that our predictions hold in an alternative setting that allows scaling of exploration R&D. Finally, we introduce patent citation behavior that overlays the R&D equilibrium. The appendix contains proofs of propositions.

2.1 Preferences and Final Good Technology

Consider the following continuous time economy. The world admits a representative household with a constant relative risk aversion utility function

$$U = \int_0^\infty \exp\left(-\rho t\right) \frac{C\left(t\right)^{1-\varepsilon} - 1}{1 - \varepsilon} dt. \tag{1}$$

$C\left(t\right)$ is consumption at time t, $\rho > 0$ is the discount rate, and $\varepsilon \geq 1$ is the constant relative risk aversion parameter (also the inverse of intertemporal elasticity of substitution).[3] The household is populated by a continuum of individuals with measure one.

Individuals consume a unique final good $Y\left(t\right)$, which is also used for R&D as discussed below. The final good is produced by labor and a continuum of intermediate goods $j \in [0,1]$ with the production technology

$$Y\left(t\right) = \frac{L^\beta\left(t\right)}{1 - \beta} \int_0^1 q_j^\beta\left(t\right) k_j^{1-\beta}\left(t\right) dj. \tag{2}$$

In this specification, $k_j\left(t\right)$ is the quantity of intermediate good j, and $q_j(t)$ is its quality. We normalize the price of the final good Y to be one in every period without loss of generality. The final good is produced competitively with input prices taken as given.

Each intermediate good j is owned by a firm f, which is run by an incumbent entrepreneur. A firm is characterized by the collection of its product lines $\mathcal{J}_f = \{j : j \text{ is owned by firm } f\}$. Similarly we will denote the product (quality) portfolio of firm f by a *multiset* $\mathbf{q}_f = \{q_j : j \in \mathcal{J}_f\}$ and denote the cardinality by n_f.[4] Each intermediate good $j \in [0,1]$ is produced with a linear technology

$$k_j = \zeta \bar{q} l_j, \tag{3}$$

where l_j is the labor input, $\bar{q} \equiv \int_0^1 q_j dj$ is the average quality in the economy, and $\zeta > 0$ is a constant. This linear specification has two implications. First, intermediate goods have the same marginal cost $w/\zeta\bar{q} > 0$, where w is the wage rate in terms of the final good. Second, the marginal product of labor in the intermediate good sector grows at the same rate as in the final good sector, generating an invariant labor allocation across sectors in steady state.[5]

---

[3] Our cross-sectional results on firms hold for any ε, but we use $\varepsilon \geq 1$ as a sufficient condition for proving the existence of a dynamic steady-state equilibrium.

[4] A multiset is a generalization of a set that can contain more than one instance of the same member.

[5] An alternative specification removes labor from the intermediate good sector and uses final goods as the only input of production.

Individuals work in four capacities: in final good production (L), in intermediate good production ($\tilde{L}$), as entrepreneurs running incumbent firms (E), or as entrepreneurs running outside firms attempting to enter the market ($\tilde{E}$). In each period, the labor market has to satisfy the following constraint

$$L + \tilde{L} + E + \tilde{E} \leq 1. \tag{4}$$

We will denote the share of entrepreneurs by $\mathbb{S} \in [0,1]$ such that $\mathbb{S} = E + \tilde{E}$. Total R&D spending is $R(t)$, and the resource constraint of the economy is $Y(t) \geq C(t) + R(t)$.

2.2   Research and Development

The last innovator in each product line owns the leading patent and has monopolist pricing power until being replaced by another firm. Intermediate producers have profit incentives to improve the technologies for their existing products, thereby increasing associated mark-ups.

In addition, both incumbents and outside entrepreneurs have incentives to add new products to their portfolios through R&D competition. The outcomes of innovation processes are stochastic in nature. A realized technology advance in product line j improves the current technology quality $q_j(t)$ by a size $s_j > 0$,

$$q_j(t + \Delta t) = (1 + s_j) q_j(t),$$

where s_j depends on both the type of R&D pursued and random realizations.

Exploitation R&D   The first type of R&D is undertaken by incumbent firms to improve their existing products. We often refer to this as *exploitation R&D* or *incremental* or *internal innovation*. To improve an existing product $j \in \mathcal{J}_f$, firm f spends

$$C_z(z_j, q_j) = c_z(z_j) q_j \tag{5}$$

units of the final good. Incremental innovations are realized with the instantaneous Poisson flow rate of $z_j \geq 0$. $c_z(\cdot) : [0, \bar{z}] \to \mathbb{R}_+$ is increasing, differentiable, strictly convex, and satisfies the following conditions for some $\bar{z} > 0$: $c_z(0) = 0$, $c_z'(0) = 0$, and $\lim_{z \to \bar{z}} c_z'(z) = \infty$. Cost (5) is proportional to the quality of the good that the firm is improving. This implies that a more advanced technology has higher R&D costs and prevents artificial scale effects (e.g., Jones 1995). When *exploitation* R&D is successful, the current quality improves by size $\lambda > 0$.

Exploration R&D   The second type of R&D is undertaken by incumbents and potential new entrants to obtain technology leadership over products not currently owned. We also refer to this as *external innovation*. Through investing

$$C_x(x, \bar{q}) = c_x(x) \bar{q} \tag{6}$$

unit of the final good, firm f innovates a new good at a Poisson flow rate $x \geq 0$. Cost (6) is proportional to the average quality level $\bar{q}$ in the economy, which again removes scale effects. The cost function $c_x(\cdot) : [0, \bar{x}] \to \mathbb{R}_+$ satisfies the same conditions as $c_z(\cdot)$.

These *exploration* R&D efforts are undirected in the sense that resulting innovations are realized in any product line $j \in [0,1]$ with equal probability. This has two main implications. First, firms do not innovate over their own product lines through exploration R&D since this event has zero probability. Second, there is no strategic interaction among firms. In addition to stochastic arrival rates, the sizes of realized quality improvements are randomly determined:

(*i*) With probability $\theta \in (0,1)$, the innovation is a *major advance* that substantially shifts forward the latest quality by a size $\eta > \lambda$. This generates a new technology cluster with an associated wave of subsequent follow-on innovations. Prominent examples include the transistor and mapping the human genome, but the step functions need not be so profound. The conceptual construct is that these major advances define a wave of innovation and product development until another major advance starts a new wave.

(*ii*) With probability $1 - \theta$, the innovation is a *follow-up improvement* to the current technology level of the product line that does not generate a new technology cluster. The size of the follow-up improvement declines with the number of follow-up inventions since the last major advancement. If the last major innovation in product line j occurred k_j innovations ago, the new step size is $s_j = \eta \alpha^{k_j}$ with $\alpha \in (0,1)$. These improvements can be larger or smaller than λ depending upon α and k_j.

Let us denote the economy-wide arrival rate of a new product by τ, which is endogenously determined by the *exploration* R&D efforts of incumbents and potential entrants. To summarize, the probabilistic evolution of the quality level q_j after a short interval Δt is

$$
q_j (t + \Delta t) = \begin{cases} (1 + \eta)\, q_j(t) & \text{with probability} & \tau \Delta t \theta \\ (1 + s_j)\, q_j(t) & \text{with probability} & \tau \Delta t (1 - \theta) \\ (1 + \lambda)\, q_j(t) & \text{with probability} & z_j \Delta t \\ q_j(t) & \text{with probability} & 1 - z_j \Delta t - \tau \Delta t \end{cases}
$$

The first line represents a major advance that results from exploration R&D with probability θ. The second line represents a follow-up innovation that results from exploration R&D with probability $1 - \theta$. The third line shows an incremental improvement of size λ by the current owner of product line j through exploitation R&D. The final line represents the case where no quality improvement is realized during Δt, which results in stagnant technology quality.

The following example illustrates a possible evolution of innovations in a random product line:

<div align="center">Example 1</div>

<div align="center">An example of a sequence of innovations in a product line</div>
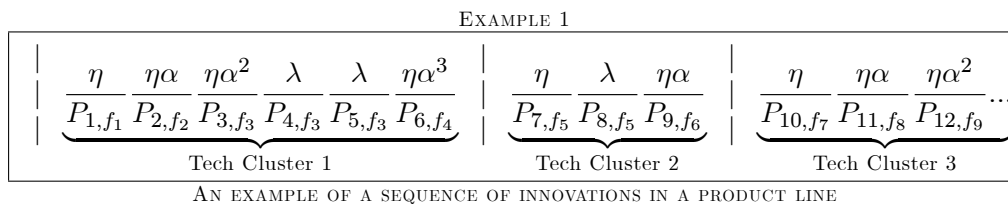
Here, $P_{m,f}$ denotes that the mth patent is obtained by firm f. The example starts with a major innovation that opens a new technology cluster by firm f_1. Firms f_2 and f_3 then produce follow-up innovations. Firm f_3 further improves its own product twice. Firm f_4 then produces a further follow-up innovation. Next, this technology cluster is replaced by a new leading innovation by firm f_5, which is patented as P_7. The second cluster is then replaced by another leading innovation by firm f_7. This new cluster is further improved by patents 11 and 12, and so on.

2.3 Entry and Exit

There are two types of entrepreneurs in the economy. *Incumbent entrepreneurs* (E) run firms that own existing product lines and invest in R&D, whereas *outside entrepreneurs* ($\tilde{E}$) invest

<div align="center">7</div>

only in R&D in order to become intermediate producers upon a successful innovation. Outside entrepreneurs choose an innovation flow rate $\nu > 0$ with an R&D cost $C_x(\nu, \bar{q}) = c_x(\nu)\bar{q}$ in terms of the final good. The value V_0 of being an outside entrepreneur is the expected value from innovating successfully and entering the market. This value is determined according to

$$rV_0 - \dot{V}_0 = \max_{\nu \in [0,\bar{x}]} \{\nu [\mathbb{E}_j V(\{q_j(1+s_j)\}) - V_0] - c_x(\nu)\bar{q}\} \tag{7}$$

where $V(\{q\})$ denotes the value of a firm that owns a single product line with quality q and $\dot{V}_0 \equiv \partial V_0/\partial t$ denotes the partial derivative of the outside value with respect to time. The expected value $\mathbb{E}_j V(\{q_j(1+s_j)\})$ of a new innovation is an expectation over both quality level q and innovation size s_j. Production workers strictly prefer to become outside entrepreneurs when the outside value V_0 exceeds the present discounted sum of the future wage stream w/r and vice versa. The *no-arbitrage condition* will pin down the number of entrepreneurs as follows

$$\mathbb{S} \begin{cases} = & 1 & \text{if} & V_0 > w/r \\ \in & [0,1] & \text{if} & V_0 = w/r \\ = & 0 & \text{if} & V_0 < w/r \end{cases} \tag{8}$$

Incumbent entrepreneurs produce intermediate inputs and invest in R&D. As a result, firms simultaneously expand into new product lines and lose some of their current product lines to other firms in the economy through competition. Each product line will face an aggregate endogenous destruction rate, which will be denoted by τ. An entrepreneur who has lost all product lines to competitors becomes an outside entrepreneur at that instant with value V_0.[6]

Definition 1 (Allocation) *In this economy, an allocation at every instant t consists of aggregate output $Y(t)$, aggregate consumption $C(t)$, total R&D spending $R(t)$, final sector production workers $L(t)$, intermediate sector production workers $\tilde{L}(t)$, incumbent entrepreneurs $E(t)$, outside entrepreneurs $\tilde{E}(t)$, price $p_j(t)$ and quantity $k_j(t)$ for each intermediate good j, R&D decisions by incumbents to develop a new product $x(t)$ and to improve each intermediate good j by $z_j(t)$, R&D decision by outsiders $\nu(t)$, wage rate $w(t)$, and interest rate $r(t)$.*

2.4 Equilibrium

We now characterize the Markov Perfect Equilibria of the economy that make strategies a function of payoff-relevant states only. We focus on the steady state in which aggregate variables $(Y, C, R, w, \bar{q})$ grow at the constant rate g. Asterisks "$*$" denote equilibrium values.

Definition 2 (Equilibrium) *In this economy, a Markov Perfect Equilibrium is a sequence*

$$\left\{ p_j^*(t), k_j^*(t), x^*(t), z_j^*(t), \nu^*(t), L^*(t), \tilde{L}^*(t), E^*(t), \tilde{E}^*(t), Y^*(t), w^*(t), r^*(t) \right\}_{t \geq 0}$$

such that (i) $p_j^*(t)$ and $k_j^*(t)$ maximize intermediate good producer's operating profits, (ii) $x^*(t)$, $z_j^*(t)$ and $\nu^*(t)$ maximize intermediate good producer's and outside entrepreneur's firm

---

[6]Many growth models do not introduce entrepreneurs. However, firm exit is a key feature in firm dynamics. In contrast to Klette and Kortum (2004), each firm in our model has an option value $V_0 > 0$ from staying in the market due to exploration R&D. As a result, a firm that loses all of its product lines does not necessarily exit the market. Introducing entrepreneurs into the model lets the entrepreneurs utilize their outside option when they lose all of their product lines and generates smooth firm exits.

value, (iii) labor allocations $L^{*}(t), \tilde{L}^{*}(t), E^{*}(t)$ and $\tilde{E}^{*}(t)$ are optimal in the sense that they maximize intermediate and final good producer's profit and also satisfy the worker's no-arbitrage condition (8) and labor market clearing condition (4), (iv) $Y^{*}(t)$ is given in (2), (v) $w^{*}(t)$ clears the labor market, and (vi) $r^{*}(t)$ is consistent with the household's intertemporal consumption choice.

Now we are ready to solve the model starting from the household's problem.

2.4.1   Production

The standard maximization problem of the representative household yields the Euler equation

$$\frac{\dot{Y}^{*}}{Y^{*}} = \frac{\dot{C}^{*}}{C^{*}} = \frac{r^{*} - \rho}{\varepsilon}. \tag{9}$$

The maximization problem of the final goods producer generates the inverse demand $p_{j} = L^{*\beta}q_{j}^{\beta}k_{j}^{-\beta}$, $\forall j \in [0,1]$. The constant marginal cost of producing each intermediate variety is $w^{*}/\zeta\bar{q}$, and the profit maximization problem of the monopolist j is

$$\pi^{*}(q_{j}) = \max_{k_{j} \geq 0} \left\{ L^{*\beta}q_{j}^{\beta}k_{j}^{1-\beta} - \frac{w^{*}}{\zeta\bar{q}}k_{j} \right\} \ \forall j \in [0,1].$$

The first order condition yields an optimal quantity and price for intermediate good j

$$k_{j}^{*} = \left[\frac{(1-\beta)\zeta\bar{q}}{w^{*}} \right]^{\frac{1}{\beta}} L^{*}q_{j} \text{ and } p_{j}^{*} = \frac{w^{*}}{(1-\beta)\zeta\bar{q}}. \tag{10}$$

The realized price is a constant markup over the marginal cost and is independent of the individual product quality. Thus, the profit for each active good is $\pi^{*}(q_{j}) = \pi^{*}q_{j}$, where $\pi^{*} \equiv L^{*}\left(\zeta\bar{q}/w^{*}\right)^{\frac{1-\beta}{\beta}}(1-\beta)^{\frac{1-\beta}{\beta}}\beta$.

The maximization in the final goods sector, together with (10), implies a wage rate

$$w^{*} = \tilde{\beta}\bar{q} \tag{11}$$

where $\tilde{\beta} \equiv \beta^{\beta}\left[1-\beta\right]^{1-2\beta}\zeta^{1-\beta}$ and $\bar{q}$ is the average quality index defined earlier. Incorporating the equilibrium wage rate, the constant part of the equilibrium profit simplifies to

$$\pi^{*} = L^{*}(1-\beta)\tilde{\beta}. \tag{12}$$

Equations (3), (10), and (11) determine the ratio of workers employed in the intermediate good sector to the final good sector

$$\frac{\tilde{L}^{*}}{L^{*}} = \frac{(1-\beta)^{2}}{\beta}. \tag{13}$$

2.4.2 Research and Development by Incumbents

The value functions of firms determine R&D choices. $\mathbf{q}_f$ denotes the product portfolio of firm f and serves as the state variable in firm's problem. Firm f takes the equilibrium values of $(r^*, \tau^*, \pi^*, g^*)$ as given and chooses the optimal R&D efforts x and z_j for every $j \in \mathcal{J}_f$ to maximize the following value function[7]

$$r^* V(\mathbf{q}_f) - \dot{V}(\mathbf{q}_f) = \max_{\substack{x \in [0, \bar{x}], \\ \{z_j \in [0, \bar{z}]\}_{\mathcal{J}_f}}} \left\{ \begin{array}{c} \sum_{q_j \in \mathbf{q}_f} \pi^* q_j - \sum_{q_j \in \mathbf{q}_f} c_z(z_j) q_j - c_x(x) \bar{q} \\ + x \left[\begin{array}{c} \theta \mathbb{E}_j V(\mathbf{q}_f \cup_+ \{q_j(1+\eta)\}) \\ + (1-\theta) \mathbb{E}_j V(\mathbf{q}_f \cup_+ \{q_j(1+s_j)\}) - V(\mathbf{q}_f) \end{array} \right] \\ + \sum_{q_j \in \mathbf{q}_f} z_j \left[V(\mathbf{q}_f \setminus_- \{q_j\} \cup_+ \{q_j(1+\lambda)\}) - V(\mathbf{q}_f) \right] \\ + \sum_{q_j \in \mathbf{q}_f} \tau^* \left[V(\mathbf{q}_f \setminus_- \{q_j\}) - V(\mathbf{q}_f) \right] \end{array} \right\} \cdot^8 \tag{14}$$

The first line on the right hand side represents operating profits over currently held product lines minus R&D costs.

The second line is the change in firm value after a successful innovation that garners a new product line. $V(\mathbf{q}_f \cup_+ \{q_j(1+s_j)\})$ denotes equilibrium firm value after a successful exploration innovation of size s_j, which adds a new product into firm's portfolio. With probability θ, exploration R&D generates a major advance. With probability $(1-\theta)$, a follow-up advance occurs. In the case of a major innovation, the step size is η and expectations are only over the quality q of the acquired product line. For follow-up advances, the expectation is also over the innovation size $s_j = \eta \alpha^{k_j}$. These terms are multiplied by the Poisson arrival rate x, as the success of *exploration* R&D is stochastic.

The third line is the change in firm value after incremental improvements to currently held products. $V(\mathbf{q}_f \setminus_- \{q_j\} \cup_+ \{q_j(1+\lambda)\})$ denotes the firm value after improving one of the firm's existing products by size λ. These terms are multiplied by the Poisson arrival rate z_j as the success of *exploitation* R&D is stochastic, too. Firms choose innovation effort for each line separately.

The fourth line shows the change in firm value due to losing its product lines through creative destruction τ^*. $V(\mathbf{q}_f \setminus_- \{q_j\})$ denotes firm value after losing product that had quality q_j. The $-\dot{V}(\mathbf{q})$ term on the left hand side of equation (14) represents change in firm value without any material events for the focal firm due to economy-wide growth (i.e., $\bar{q}$ changes).

The next proposition shows that the value function (14) can be expressed in a very tractable form that is central for the rest of the analysis. For notational simplicity, we define the expected innovation size $\Gamma \equiv \mathbb{E}_j(s_j)$.

Proposition 1 *For given equilibrium values of $(r^*, \tau^*, \pi^*, g^*)$, the value function (14) of a firm with a set of product lines $\mathbf{q}_f$ can be expressed as*

$$V(\mathbf{q}_f) = A \sum_{q \in \mathbf{q}_f} q + B \bar{q} \tag{15}$$

---

[7]We do not index the R&D efforts by f as x_f and $z_{j,f}$ to simplify notation.

[8]$\cup_+$ indicates the multiset union operator such that $\{a, b\} \cup_+ \{b\} = \{a, b, b\}$. Similarly $\setminus_-$ indicates the multiset difference operator such that $\{a, b, b\} \setminus_- \{b\} = \{a, b\}$.

where A (value of holding a product line) and B (value of innovating a new product line) are defined by

$$(r^* + \tau^*) A = \pi^* + \max_{z \in [0, \bar{z}]} \{A\lambda z - c_z (z)\} \tag{16}$$

and

$$(r^* - g^*) B = \max_{x \in [0, \bar{x}]} \{Ax (1 + \Gamma) - c_x (x)\}, \tag{17}$$

where g^* is the equilibrium growth rate of the average quality in the economy. The optimal R&D decisions are given by

$$z^* = c_z'^{-1} (A\lambda) \ \ and \ x^* = c_x'^{-1} (A [1 + \Gamma]). \tag{18}$$

The proposition demonstrates that the exploitation R&D spending on incremental innovations by incumbents is proportional to the number of product lines that the incumbent owns. Proposition 1 further shows that R&D decisions are independent of a firm's quality portfolio $\mathbf{q}_f$. Finally, the proposition implies that the outside entrepreneur's value is

$$V_0 = B\bar{q}. \tag{19}$$

2.4.3 Research and Development by Outsiders, Entry and Exit

The first order condition of the optimization problem (7), together with (15) and (19), determine the optimal R&D efforts by outsider entrepreneurs ν^*,

$$\nu^* = c_x'^{-1} (A [1 + \Gamma]) = x^*, \tag{20}$$

where the last equality follows from (18). These R&D efforts yield an equilibrium creative destruction rate of

$$\tau^* = E^* x^* + \tilde{E}^* \nu^* = \mathbb{S}^* x^*, \tag{21}$$

where we sum new product arrivals by the incumbents and outside entrepreneurs.

A common feature of this class of models is that an equilibrium can exist with no entry $(\mathbb{S} = 0)$. To focus on an empirically relevant equilibrium with positive entry, we make the following assumption with respect to parameter values.

Assumption 1 (Positive entry) *We assume that the parameters of the model satisfy*

$$\max_{x \in [0, \bar{x}]} \left\{ x \frac{\beta (1 - \beta)}{[1 - \beta (1 - \beta)]} \frac{(1 + \Gamma)}{(\lambda \bar{z} (\varepsilon - 1) + \rho)} - \frac{c_x (x)}{\beta^\beta [1 - \beta]^{1 - 2\beta} \zeta^{1 - \beta}} \right\} > 1.$$

This condition is formally derived in the proof of Proposition 4. It requires that the expected value of being an entrepreneur is more profitable than being a worker when all members of the economy are workers. This ensures an interior solution with positive entry rates. Intuitively, for any given β, this assumption is satisfied when the expected arrival rate x or step size Γ from exploration R&D are sufficiently high relative to costs $c_x (x)$. The condition is also satisfied for sufficiently high labor productivity ζ or a small enough discount factor ρ.

On the other hand, the value of an outside entrepreneur is smaller than the value of being a worker when all individuals switch to being an entrepreneur (see appendix for further details). As a result, the no-arbitrage condition (8) holds in equilibrium with equality,

$$B^* = \tilde{\beta}/r^*, \tag{22}$$

where we combined (8) together with (11) and (19).

11

2.4.4 Invariant Distributions and Labor Demand

Next, we determine the expected innovation size Γ and the measure of outside entrepreneurs. To accomplish this task, we need to solve for two invariant distributions: innovation size distribution (Ψ) and number of product distribution (μ). We now characterize the innovation size distribution. This distribution determines the expected innovation size Γ. Let Ψ_k^* denote the equilibrium share of product lines with k subsequent follow-up innovations. A steady state equilibrium requires a stable innovation size distribution. Thus, while stochastic innovation properties move individual products up and down the k distribution, the overall share of products at each k is stable. This stability requires equal inflows and outflows of products from each size level, resulting in the flow equations

$$
\begin{array}{ccc}
\text{OUTFLOW} & & \text{INFLOW} \\
\Psi_0^* \tau^* (1 - \theta) & = & (1 - \Psi_0^*) \tau^* \theta \\
\Psi_k^* \tau^* & = & \Psi_{k-1}^* \tau^* (1 - \theta) \ \text{ for } k \geq 1.
\end{array}
$$
(23)
(24)

The first line governs inflows and outflows among product lines where major innovations have just occurred. Outflows happen due to follow-up innovations at the rate $\tau^* (1 - \theta)$, while inflows happen due to new leading innovations being realized at rate $\tau^* \theta$ throughout the innovation size distribution. By assumption, incremental innovations within firms do not influence these k distributions. A similar reasoning governs the share of product lines with $k \geq 1$ consecutive follow-up innovations in (24). As a result, flow equations (23) and (24) generate the invariant distribution

$$
\Psi_k^* = \theta (1 - \theta)^k \ \text{ for } k \geq 0,
$$
(25)

which yield the expected innovation size $\Gamma = \theta\eta + (1 - \theta) \sum_{k=0}^{\infty} \Psi_k^* \eta \alpha^{k+1} = \theta\eta / (1 - (1 - \theta) \alpha)$.

We next characterize the invariant distribution of the number of products. This distribution is the main proxy for the firm size distribution in Klette and Kortum (2004) and Lentz and Mortensen (2008). Let μ_n^* denote the equilibrium share of the incumbent firms that own n product lines. The invariant distribution again depends upon flow equations

$$
\begin{array}{lccc}
\text{STATE :} & \text{INFLOW} & & \text{OUTFLOW} \\
n = 0 : & E^* \mu_1^* \tau^* & = & \tilde{E}^* \nu^* \\
n = 1 : & E^* \mu_2^* 2\tau^* + \tilde{E}^* \nu^* & = & E^* \mu_1^* (x^* + \tau^*) \\
n \geq 2 : & E^* \mu_{n+1}^* (n + 1) \tau^* + E^* \mu_{n-1}^* x^* & = & E^* \mu_n^* (x^* + n\tau^*)
\end{array}
$$
(26)

The first line characterizes outside entrepreneurs ($n = 0$). Inflows to outside entrepreneurs happen when firms with one product are destroyed, and outflows occur when outside entrepreneurs successfully develop a new product at rate ν^*. Similarly, the second line considers inflows and outflows of firms with one product, and the third line considers n-product firms. The next proposition provides the explicit form solution of the invariant product number distribution.

Proposition 2 *The invariant distribution μ_n^* is equal to*

$$
\mu_n^* = \frac{\tilde{E}^*}{E^*} \left(\frac{x^*}{\tau^*} \right)^n \frac{1}{n!} \ for \ n \geq 1.
$$
(27)

As firms can own multiple product lines, the following identity on the number of product lines holds, $\sum_{n=1}^{\infty} E^* \mu_n^* n = 1$. Using this fact, together with (21) and (27), the total measure of outside entrepreneurs satisfies

$$\tilde{E}^* = \mathbb{S}^* \exp\left(-\mathbb{S}^{*-1}\right). \tag{28}$$

Labor market clearing requires that (4) holds with equality,

$$L^* + \tilde{L}^* + E^* + \tilde{E}^* = 1. \tag{29}$$

This result, together with (13), (22), and (28) pin down all labor allocations. Next, we express the equilibrium growth rate of the economy.

Proposition 3 *Let the equilibrium R&D efforts be given by (τ^*, z^*). The steady state growth rate of the aggregate variables in the economy is*

$$g^* = \tau^* \Gamma + z^* \lambda \tag{30}$$

Now we are ready to summarize the equilibrium as follows.

Definition 3 (Balanced Growth Path Equilibrium) *A balanced growth path equilibrium of this economy consists of the following tuple for every t, $j \in [0,1]$ and q_j*

$$\{Y^*(t), w^*(t), k_j^*(t), p_j^*(t), x^*, z_j^*, \nu^*, L^*, \tilde{L}^*, E^*, \tilde{E}^*, g^*, \Psi_n^*, \mu_n^*, r^*\}$$

such that: (i) $Y^*(t)$ and $w^*(t)$ satisfy (9) and (11); (ii) $k_j^*(t)$ and $p_j^*(t)$ satisfy (10); (iii) x^* and z_j^* solve the value function in (14); (iv) ν^* satisfies the FOC of the free entry condition in (20); (v) labor measures $\tilde{L}^*, \tilde{E}^*, E^*$, and L^* satisfy (13), (22), (28), and (29); (vi) the invariant distribution of innovation sizes satisfies (25); (vii) steady state equilibrium growth rate satisfies (30); (viii) the invariant distribution of number of products satisfies (26); and (ix) the equilibrium interest rate satisfies the Euler equation (9).

Using this definition, one can reduce equilibrium further into a triplet $(x^*, z^*, \mathbb{S}^*)$ and define a continuous correspondence Λ that maps the compact set $\Theta \equiv [0, \bar{x}] \times [0, \bar{z}] \times [0, 1]$ into itself $\Lambda : \Theta \rightrightarrows \Theta$. Existence of a fixed point of this mapping is guaranteed by Brouwer's fixed point theorem. Then the equilibrium of the model is simply the fixed point of this mapping

$$\Lambda\left([x^* \quad z^* \quad \mathbb{S}^*]^T\right) = [x^* \quad z^* \quad \mathbb{S}^*]^T.$$

The details of this mapping are provided in the proof of the following proposition.

Proposition 4 *There exists a steady state equilibrium with strictly positive growth $g^* > 0$ and strictly positive entry $\mathbb{S} \in (0,1)$.*

13

2.5 Central Theoretical Results

The following propositions characterize the firm growth, R&D, and innovation dynamics of the model, which we elaborate on in Section 2.8. We initially proxy firm size by $Q_f = \sum_{q_j \in \mathbf{q}_f} q_j$. Firm sales, profits, and production workers are proportional to Q_f.[9] Firm size also closely relates to the number of product lines.

Lemma 1 *Expected firm size is strictly increasing in the number of products*

$$\mathbb{E}\left(Q_f \mid n_f\right) > \mathbb{E}\left(Q_{f'} \mid n_{f'}\right) \text{ for any } n_f > n_{f'}.$$

Therefore, when convenient, we also use n_f to proxy for firm size in propositions.

Proposition 5 *Let $\mathcal{G}\left(Q_f\right) \equiv \mathbb{E}\left(\dot{Q}_f / Q_f\right)$ be the average growth rate of a firm with total quality Q_f. Then $\mathcal{G}\left(Q_f\right)$, in equilibrium, is given by*

$$\mathcal{G}\left(Q_f\right) = \frac{x^*\left(1 + \Gamma\right)\bar{q}}{Q_f} + z^*\lambda - \tau^*.$$

$\mathcal{G}\left(Q_f\right)$ *is a strictly decreasing function.*

This result suggests that small firms grow faster than large firms. This micro-founded departure from Gibrat's Law of proportionate growth is one of the central results of our work. The intuition behind this result is as follows. Exploitation innovation scales up with firm size, yet exploration innovation does not. As a result, the growth coming from exploitation innovation is the same on average across different firm sizes $(z^*\lambda)$, whereas the contribution of exploration R&D to firm growth gets smaller as firm size increases (the first ratio in $\mathcal{G}\left(Q_f\right)$). Combining these effects, overall firm growth declines with firm size.

Proposition 6 *Let $\mathcal{R}\left(Q_f\right) \equiv R\&D/Sales$ be the firm R&D intensity of a firm with total quality Q_f. Then $\mathcal{R}\left(Q_f\right)$, in equilibrium, is given by*

$$\mathcal{R}\left(Q_f\right) = \frac{\beta c_x\left(x^*\right)\bar{q}}{\pi^* Q_f} + \frac{\beta c_z\left(z^*\right)}{\pi^*}.$$

$\mathcal{R}\left(Q_f\right)$ *is a strictly decreasing function.*

This result suggests that small firms have a greater R&D intensity than large firms. Similar to the previous proposition, the intuition is that total exploitation R&D effort is proportionate to the number of product lines of the firm. On the other hand, exploration R&D efforts do not scale with number of product lines, which results in a declining R&D intensity for larger firms. In other words, adding additional product lines continually adds more R&D efforts but further dilutes the exploration R&D effects with respect to intensity measures.

---

[9] $Sales = \sum_{q_j \in \mathbf{q}_f} p\left(q_j\right) k\left(q_j\right) = \left[(1 - \beta)/\psi\right]^{\frac{1-\beta}{\beta}} L^* Q_f$, $Profits = \sum_{q_j \in \mathbf{q}_f} \pi^* q_j = \pi^* Q_f$, and $Production$ $workers = \sum_{q_j \in \mathbf{q}_f} l_j = \left[(1 - \beta)/w^*\right]^{\frac{1}{\beta}} \zeta^{\frac{1-\beta}{\beta}} L^* Q_f$.

Proposition 7 *Let a major innovation be defined as an innovation with a step size larger than a certain threshold $s_k \geq s_{\hat{k}}$ for some $\hat{k} \in \mathbb{Z}_+$ and $s_{\hat{k}} > \lambda$. Moreover, let $\mathcal{M}(n_f)$ be the probability of making a major innovation conditional on having a successful innovation for a firm with n_f product lines. Then, $\mathcal{M}(n_f)$ can be expressed by*

$$\mathcal{M}(n_f) \equiv \frac{x^* \sum_{k=0}^{\hat{k}} \theta (1-\theta)^k}{x^* + n_f z^*} = \frac{x^* \left[1 - (1-\theta)^{\hat{k}+1}\right]}{x^* + n_f z^*}.$$

$\mathcal{M}(n_f)$ *is a strictly decreasing function.*

This result suggests that small firms and new entrants have a comparative advantage for achieving major advances. Large incumbents endogenously spend effort on maintaining and expanding existing products. Thus, while firms of all sizes obtain major advances, these major advances account for a smaller share of achieved innovations among larger firms. An important distributional implication of Proposition 7 is that these differences weaken when considering progressively larger thresholds $s_{\hat{k}}$. The comparative advantage is weakest at the most extreme values (i.e., $s_{\hat{k}=0} = \eta$).[10]

2.6   Discussion of the Baseline Model

Our theoretical framework provides a very tractable environment to study the heterogenous innovation behavior of different sized firms. We next review our main assumptions and demonstrate that our results hold in alternative frameworks.

2.6.1   Key Model Assumptions

Assumption 2 (Structure of step sizes) *Within a technology cluster, exploration innovation sizes $s_j = \eta \alpha^{k_j}$ decrease with the number of steps since the major invention. On the other hand, exploitation innovation improves quality by a constant λ.*

The first part of this assumption is important and is motivated by the fact that some innovations are radical in nature and substantially shift the evolution of product lines and industries. Similarly, major breakthroughs can allow for important follow-on innovations. We assume a decreasing return over time unless another major innovation wave emerges to reflect diminishing productivity as easily identifiable ideas are exhausted. This structure also allows for our upcoming connection between patent citations and underlying innovations.

The second assumption of a constant exploitation innovation size simplifies the analysis, but the core results of the model could instead build upon incumbents drawing stochastically from a distribution of exploitation innovation sizes. To incorporate the second distribution, one must keep track of the two invariant distributions of exploration and exploitation innovation sizes and their joint distribution. This substantially complicates the model without generating further insights.

Assumption 3 (Common research technology) *The realized outcomes of types of innovation do not depend upon firm traits.*

---

[10] The aggregate quantity of major innovations by small and large firms depends upon these propensities and the firm size distribution.

Our framework assumes that successful innovative efforts of small and large firms draw from the same distribution of outcomes for each type of R&D pursued. In other words, research capabilities are the same for all firms, but firms endogenously undertake different types of R&D. This is a very important baseline that we verify empirically in Figures 4 and 5 below.

Assumption 4 (Scaling up of R&D) *Exploration R&D does not scale with firm size compared to exploitation R&D.*

This concept is central to our model and is embedded in the single stream of exploration R&D per firm. Figure 1 and our extended analysis in Section 3.5 characterize this difference. The specific formulation that exploration R&D does not scale at all, however, is a simplification that improves tractability. The true requirement is that exploration R&D not scale as quickly with firm size as refinement R&D does.

As a more formal description, and to prepare for the next extension, define X_n to be the total instantaneous arrival rate of exploration innovations in a firm with n existing products. Likewise, $x_n \equiv X_n/n$ represents innovations per product line or innovation intensity. In our benchmark model, all firms choose the same optimal total exploration R&D effort, such that $X_n^* = X^*, \forall n$. Our model and core results, however, generalize to any setting where exploration innovation intensity x_n^* is decreasing in firm size. R&D technologies that allow for partial scaling of X_n with firm size are compatible with our conclusions so long as the intensity (X_n^*/n) remains decreasing in n. We next consider a case that allows for scaling of exploration R&D.

2.6.2 Klette and Kortum (2004) R&D Technology

In their seminal study, Klette and Kortum (2004) focus on exploration R&D, which they allow to scale with firm size, and do not consider exploitation R&D to improve currently held product lines. They model the R&D cost function as

$$C_k \left(X_n, n \right) = n c_k \left(\frac{X_n}{n} \right), \tag{31}$$

where X_n is defined above. $c_k \left(\cdot \right)$ is a strictly convex, differentiable cost function. The division of X_n by n in specification (31) results in marginal costs for exploration R&D increasing at the product level, as opposed to the firm level. Thus, the intensity of innovations per product line x_n is independent of firm size, which falls outside of our generalization above.

Can one substitute the exploration R&D technology structure of Klette and Kortum into our benchmark model? The answer is yes, but additional structure is required to again have the heterogeneity in innovations that are the focus of this paper and our empirical observations. To keep the analysis tractable, we will focus on the maximization of a single firm, abstracting from the aggregate innovation in the economy, so that $g = \tau = 0$ and $\bar{q} > 0$ is fixed. If one only substitutes (31) for (6), incumbent firms undertake equivalent amounts of exploration and exploitation R&D per product line in expectation. Total R&D effort exactly scales with firm size, as does its mix. Each firm would identically decide for the same exploration and exploitation R&D efforts at the product line level. The resulting model no longer contains any frictions that are capable of explaining our findings in the data.[11]

---

[11] In most respects, the model becomes equivalent to single R&D technology for incumbent firms that both

16

Yet, the model becomes aligned with the data's heterogeneity again if we also introduce a cost for holding more product lines

$$C_n(n) = nc_m(n), \tag{32}$$

where $c_m'(n) > 0$, $c_m''(n) > 0$ and $c_m(0) = 0$. This cost function simply models that the cost of managerial efforts to coordinate more production units/divisions grows as the number of products increases.

These coordination costs and associated frictions allow the augmented model to generate interesting heterogeneities again. Managerial and coordination costs re-introduce heterogeneity between exploration and exploitation investments even with the exploration cost function (31). Exploration efforts lead to growth in the number of product lines, while exploitation efforts do not. Thus, in the presence of these coordination costs, smaller firms have a comparative advantage for exploration investments because product line expansion is less costly for them. Larger firms, with more product lines but higher management costs, have a comparative advantage for exploitation.

Formally, we rewrite the main value function (14) as

$$r^*V(\mathbf{q}_f) - \dot{V}(\mathbf{q}_f) = \max_{\substack{x_n \in [0,\bar{x}], \\ \{z_j \in [0,\bar{z}]\}_{\mathcal{J}_f}}} \left\{ \begin{array}{c} \sum_{q_j \in \mathbf{q}_f} [\pi^* q_j - c_z(z_j) q_j] - nc_k(x_n) - nc_m(n) \\ +nx_n \left[\begin{array}{c} \theta \mathbb{E}_j V(\mathbf{q}_f \cup_+ \{q_j(1+\eta)\}) \\ + (1-\theta) \mathbb{E}_j V(\mathbf{q}_f \cup_+ \{q_j(1+s_j)\}) - V(\mathbf{q}_f) \end{array} \right] \\ + \sum_{q_j \in \mathbf{q}_f} z_j [V(\mathbf{q}_f \setminus \{q_j\} \cup_+ \{q_j(1+\lambda)\}) - V(\mathbf{q}_f)] \end{array} \right\}. \tag{33}$$

The next proposition characterizes this value function and its key properties.

Proposition 8 *The value function* (33) *of an* n_f-*product firm with a set of product lines* $\mathbf{q}_f$ *can be expressed as*

$$V(\mathbf{q}_f) = \tilde{A} \sum_{q \in \mathbf{q}_f} q + W(n_f) \tag{34}$$

where $\tilde{A}$ *(value of holding a product line) is defined by*

$$r^*\tilde{A} = \max_{z \in [0,\bar{z}]} \left\{ \pi^* - c_z(z) + \tilde{A}\lambda z \right\}$$

and $W(n)$ *(value of innovating a new product line) is defined by*

$$r^*W(n) = \max_{x_n \in [0,\bar{x}]} \{-nc_k(x_n) - nc_m(n) + nx_n[\Upsilon + W(n+1) - W(n)]\}$$

where $\Upsilon \equiv \tilde{A}(1+\Gamma)\bar{q}$. *The optimal exploration and exploitation R&D decision are given by*

$$x_n^* = c_m'^{-1}(\max\{0, \Upsilon + W(n+1) - W(n)\}) \ \ and \ z^* = c_z'^{-1}(\tilde{A}\lambda). \tag{35}$$

Moreover,

(i) $\{W(n)\}_{n=1}^{\infty}$ *is bounded from above,*

(ii) $\{x_n^*\}_{n=1}^{\infty}$ *forms a bounded, decreasing sequence such that* $x_{n+1}^* < x_n^*$ *when* $x_n^* > 0$, *and* $x_{n+1}^* = 0$ *otherwise.*

---

stochastically improves existing product qualities and discovers new technologies. The resulting structure does yield a special role for entrants, who can only undertake exploration efforts by definition, but otherwise the model yields few predictions for firm R&D heterogeneity that we study in the data.

Proposition 8 indicates that intensities of exploration R&D efforts $\{x_n^*\}_{n=1}^{\infty}$ decrease with firm size in this alternative framework. Total exploration R&D efforts $(X_n^* \equiv nx_n^*)$ can increase with larger firm size, but the intensity is always decreasing. Exploitation efforts, on the other hand, continue to scale directly with firm size. Rising management and coordination costs endogenously lead firms to focus more on their existing product lines rather than search for new ones even when R&D cost structures have similar properties. Therefore we have the following main results as before.

Proposition 9 *Consider the benchmark model's variation in Section 2.6.2. For any given firm f with $n > 0$ and the equilibrium R&D efforts $\{x_n^*\}_{n=1}^{\infty}$ and z^*,*
(i) average growth rate of the number of product lines $\mathcal{G}_n \equiv \mathbb{E}(\dot{n}/n) = x_n^*$ decreases in n,
(ii) average growth rate of sales in the existing product lines $\mathcal{G}_q \equiv z^* \lambda$ is constant across firms,
(iii) R&D intensity $\mathcal{R}_n \equiv \beta \left(c_m \left(x_n^* \right) \bar{q} + c_z \left(z^* \right) \right) / \pi^*$ decreases in n,
(iv) conditional on a successful innovation, the probability of a major innovation $(s_k \geq s_{\hat{k}} > \lambda)$
$\mathcal{M}_n \equiv x_n^* \left[1 - (1 - \theta)^{\hat{k}+1} \right] / (x_n^* + nz^*)$ *decreases in n.*

2.7 Patent Citations Behavior and Innovation Spillover Sizes

We now incorporate patent citation behavior across innovations into our benchmark model. As we have already defined the economy's equilibrium, our specified citation behavior does not affect real outcomes. We undertake this extension, however, to derive the economic meaning behind patent citations. This in turn allows us to verify our modelling assumptions. Second, this addition demonstrates how this class of endogenous growth models captures many important features from empirical literature on patent counts and citations. Constructing this link between these literatures is the central purpose of this section. Finally, this extension could provide a useful theoretical framework for future studies on the impact of policies such as intellectual property rights for innovation qualities and growth.

2.7.1 Forward Patent Citations

Innovations are clustered in terms of their technological relevances. Major innovations generate new technology clusters that last until they are overtaken by a subsequent major innovation. An example of the sequential innovation process was illustrated in Example 1 in Section 2.2.

Let $m(j, t)$ be the number of patents in the active technology cluster in product line j. For instance, if t is between the innovation times of P_3 and P_4 in Example 1, then $m(j, t) = 3$, or if t is between P_{11} and P_{12}, then $m(j, t) = 2$. Therefore the number of citable patents in the economy at time t is $M(t) = \int_0^1 m(j, t)\, dj$.

We next describe the citations distributions of patents. We specify citation behavior with a few simple rules that build upon the patent literature. Patents cite previous patents within the same technology cluster to specify how they build upon the prior work and the boundaries of the innovations. Each new patent, by definition, improves the previous technologically-relevant innovations on some dimensions. However, not all subsequent innovations improve an existing technology in the same direction. Therefore major patents with broader scope are more likely to be cited by subsequent follow-on work (e.g., Lerner 1994). We proxy this patent scope by the step size $s \in \left\{ \lambda, \eta\alpha^k \mid k \in \mathbb{N}_0 \right\}$ in our model. We assume that an innovation with size s will receive a citation from a subsequent patent within the same technology cluster with probability

18

$s\gamma$ where $\gamma \in (0, 1/\eta)$. Finally a major innovation replaces the previous cluster. Thereafter, future citations begin with the new major innovation.

Thus, the citation behavior in Example 1 would be

Cited	with prob.	Citing	Cited	with prob.	Citing
$P_1:$	$\gamma\eta$	$P_2 - P_6$	$P_6:$	$\gamma\eta\alpha^3$	$none$
$P_2:$	$\gamma\eta\alpha$	$P_3 - P_6$	$P_7:$	$\gamma\eta$	$P_8, \ P_9$
$P_3:$	$\gamma\eta\alpha^2$	$P_4 - P_6$	$P_8:$	$\gamma\lambda$	P_9
$P_4:$	$\gamma\lambda$	$P_5, \ P_6$	$P_9:$	$\gamma\eta\alpha$	$none$
$P_5:$	$\gamma\lambda$	P_6	$P_{10}:$	$\gamma\eta$	$P_{11}, P_{12} \ ...$

In reality, firms do not patent all of their internal improvements, and we assume that only $\xi \in (0, 1)$ share of these innovations are patented. This propensity to patent does not influence our primary equilibrium analysis above, but its allows for reporting differences across types.[12]

2.7.2 Invariant Distributions

With these simple modeling assumptions, we can characterize the flow properties of citations behavior. These traits depend upon the real side of the economy and provide a richer description of it. Similar to our earlier expressions, the equilibrium of the economy requires an invariant citation distribution. Let $\Upsilon_{s_k,n}$ and $\Upsilon_{\lambda,n}$ denote respectively the share of patents that are of size $\eta\alpha^k$ and λ, respectively, and receive n citations such that $\Sigma_{n=0}^{\infty}\Upsilon_{\lambda,n} + \Sigma_{k=0}^{\infty}\Sigma_{n=0}^{\infty}\Upsilon_{s_k,n} = 1$. For any given innovation size $s_k = \eta\alpha^k$, the flow equations for exploration patents with n citations take the following form

$$\text{OUTFLOW} \qquad\qquad \text{INFLOW}$$
$$M^*\Upsilon_{s_k,0}\tau^*\theta + M^*\Upsilon_{s_k,0}\gamma\eta\alpha^k\left(\tau^*\left(1-\theta\right) + z^*\xi\right) = \Psi_{k-1}\tau^*\left(1-\theta\right) \ \text{for} \ n = 0. \qquad (36)$$
$$M^*\Upsilon_{s_k,n}\tau^*\theta + M^*\Upsilon_{s_k,n}\gamma\eta\alpha^k\left(\tau^*\left(1-\theta\right) + z^*\xi\right) = M^*\Upsilon_{s_k,n-1}\gamma\eta\alpha^k\left(\tau^*\left(1-\theta\right) + z^*\xi\right) \ \text{for} \ n \in \mathbb{Z}_{++}. \quad (37)$$

The first line represents size s_k innovations with no citations ($n = 0$). There are $M^*\Upsilon_{s_k,0}$ such patents for each innovation size s_k. The first part of the outflow occurs when the technology cluster is replaced through a new major innovation at the rate $\tau^*\theta$. When this happens, patents become defunct and are no longer considered for citation. The second part of the outflow occurs when patents receive a new citation from subsequent innovations at the rate $\gamma\eta\alpha^k\left(\tau^*\left(1-\theta\right) + z^*\xi\right)$. This latter expression is the probability of citation based on step size of $\gamma\eta\alpha^k$ multiplied by the arrival rate of subsequent patents. In this case, patents remain active but move up the citation distribution.

On the right hand side, the inflow occurs from Ψ_{k-1} product lines where the latest follow-up innovation was of size $\eta\alpha^{k-1}$ and a new follow-up innovation brings the product line into the Ψ_k group. This occurs at rate $\tau^*\left(1-\theta\right)$. This inflow is not dependent on the number of citable patents M^*. All patents initially have zero citations, and only a single patent can

---

[12]Caballero and Jaffe (1993) connect an endogenous growth model with patent citations to quantify the potency of new ideas, their diffusion rates, and their obsolescence rates. Their model allows them to characterize annual shifts in the aggregate rates of these variables for the US economy since the 1960s via patent citations. This antecedent relates to the aggregate creative destruction rates in our model. Our current framework further considers the heterogeneity that exists in the economy across innovations and firms as evident in patent citations.

arrive per product line at any instant. The inflow thus depends on the product lines affected only.

Similar reasoning applies to the second row, where citations $n \geq 1$, except that the inflow occurs only from the $(k, n-1)$ group. These innovations arrive at rate $\tau^* (1 - \theta) + z^*\xi$, and they cite the specific patent at rate $\gamma \eta \alpha^k$.

Next we characterize the citation distribution of exploitation patents with flow equations

$$\text{OUTFLOW} \qquad \text{INFLOW}$$

$$M^*\Upsilon_{\lambda,0}\tau^*\theta + M^*\Upsilon_{\lambda,0}\gamma\lambda\left(\tau^*\left(1-\theta\right)+z^*\xi\right) = z^*\xi \text{ for } n = 0 \tag{38}$$

$$M^*\Upsilon_{\lambda,n}\tau^*\theta + M^*\Upsilon_{\lambda,n}\gamma\lambda\left(\tau^*\left(1-\theta\right)+z^*\xi\right) = M^*\Upsilon_{\lambda,n-1}\gamma\lambda\left(\tau^*\left(1-\theta\right)+z^*\xi\right) \text{ for } n \in \mathbb{Z}_+ \tag{39}$$

These flows have similar interpretation. The substantive difference is that the inflow of zero-cited patents occurs at rate $z^*\xi$ for internal improvements. The next proposition provides the explicit form solutions for these distributions.

Proposition 10 *The invariant distribution of the total number of forward citations (n) given to a patent of size $s \in \{\lambda, s_k \mid k \in \mathbb{N}_0\}$ can be expressed as*

$$\Upsilon_{s,n} = \Upsilon_{s,0}\Omega_s^n \text{ for } n \in \mathbb{N}_0,$$

where $M^* = \frac{x^* + \xi z^*}{x^*\theta}$, $\Upsilon_{s_k,0} = \frac{\theta(1-\theta)^k\tau^*}{M^*[\tau^*\theta + \gamma s_k(\tau^*(1-\theta)+z^*\xi)]}$, $\Upsilon_{\lambda,0} = \frac{z^*\xi}{M^*[\tau^*\theta + \gamma\lambda(\tau^*(1-\theta)+z^*\xi)]}$ *and* $\Omega_s \equiv \frac{\gamma s(\tau^*(1-\theta)+z^*\xi)}{\tau^*\theta + \gamma s(\tau^*(1-\theta)+z^*\xi)}$.

Similarly, the invariant distribution of the total number of external forward citations is

$$\tilde{\Upsilon}_{s,n} = \tilde{\Upsilon}_{s,0}\tilde{\Omega}_s^n \text{ for } n \in \mathbb{N}_0,$$

where $\tilde{\Upsilon}_{s_k,0} = \frac{\theta(1-\theta)^k\tau^*}{M^*[\tau^*\theta + \gamma s_k\tau^*(1-\theta)]}$, $\tilde{\Upsilon}_{\lambda,0} = \frac{z^*\xi}{M^*[\tau^*\theta + \gamma\lambda\tau^*(1-\theta)]}$ *and* $\tilde{\Omega}_s \equiv \frac{\gamma s\tau^*(1-\theta)}{\tau^*\theta + \gamma s\tau^*(1-\theta)}$.

Note that $\Upsilon_{s,n}$ generates a more highly skewed distribution of citations as the share $\tau^*\theta$ gets smaller in the denominator. This is intuitive given the slower arrival of new technology clusters in favor of follow-on inventions that cite prior work.

2.7.3 Relative Sizes of Exploration versus Exploitation

Our last step is to connect average patent citation behavior with the underlying innovation sizes of firms. Our base model is flexible with respect to whether exploration or exploitation R&D provides stronger spillover effects into the growth process. We require that relative innovation sizes $\eta\alpha^k$ for exploration decline within clusters for follow-on innovations, but these step sizes can be larger or smaller than the λ step sizes for exploitation efforts depending upon cluster age k and the model's parameters. The described firm dynamics and equilibrium conditions above do not require further specification.

Under the conditions of our model, citation patterns have the power to discern whether exploration or exploitation innovations have a larger growth spillover effect to following entrants. Γ represents the average exploration step size, while λ is the step size for exploitation.

Lemma 2 *Average citations for exploration patents are higher than for exploitation patents if and only if $\Gamma > \lambda$.*

This lemma captures the idea that we can use patent citations to discern relative innovation sizes. Our model of citation behavior allows for stochastic features and incomplete reporting of innovations that exist in real life, but an aggregate connection between the average quality of innovations and average numbers of citations persists across types.

Recalling that $\Gamma = \eta\theta/(1 - \alpha(1 - \theta))$, we capture the intuitive conditions that promote greater spillover effects from exploration: higher step sizes η to major advances, higher likelihood θ that exploration yields a new technology cluster, and greater retention α in step sizes within technology clusters. Each of these factors raises the potential for exploration efforts relative to a given exploitation step size of λ.

Evidence presented in Section 3.4 suggests that $\Gamma > \lambda$ for the US economy as a whole during the period that we study. This suggests that the growth spillover benefits to following entrants from exploration R&D are on average higher than from exploitation R&D.

2.8 Main Predictions of the Model

We finally highlight the main testable predictions of the model, with the appendix providing formal derivations. We group predictions into five broad categories to ease the comparison between theory and empirics. Predictions A1-A5 are developed in prior work, while the remaining predictions are mostly unique to our model. Most predictions are general and do not depend upon which part of Lemma 2 holds. Prediction C3 is the outcome of the test of Lemma 2, and predictions D5 and D6 are specific to the case of $\Gamma > \lambda$.

2.8.1 Firm Size Distribution and Firm Growth Rates

A1 The size distribution of firms is highly skewed.

A2 The probability of a firm's survival is negatively related to its size.

A3 Small firms that survive tend to grow faster than larger firms. Among larger firms, this negative relationship weakens.

A4 The variance of growth rates is higher for smaller firms.

A5 Younger firms have a higher probability of exiting, but those that survive tend to grow faster than older firms.

2.8.2 Firm Size Distribution and Innovation Intensity

B1 R&D expenditures increase with firm size among innovative firms, but the intensity of R&D decreases with firm size.

B2 Similarly, patent counts increase with firm size among innovative firms, but the intensity of patenting decreases with firm size.

B3 Younger firms are more R&D and patent intensive than older firms.

2.8.3 Patent Citation Behavior and Innovation Spillover Size

C1 A large fraction of patents receive zero external citations.

C2 The distribution of citations is highly skewed.

C3 An average exploration patent receives more external citations than an exploitation patent.

C4 The distribution of patent citation life is highly skewed.

2.8.4 Innovation Type and Firm Size Distribution

D1 The proportion of a firm's patents that receives zero future external citations rises with firm size.

D2 The proportion of a firm's backwards citations that are self citations rises with contemporaneous firm size.

D3 Average future external citations per patent is decreasing in firm size.

D4 The relative rate of major innovations (highly cited patents) is higher for small firms. This higher relative rate weakens with more stringent citation quality thresholds.

D5 The average citations (received) of patents by entrants is higher than the average citations of patents by incumbents. Similarly, the average citations of patents by young firms is higher than the average citations of patents by older firms.

D6 The patents made by firms at their entry on average receive more external citations than later patents of the same firm.

2.8.5 Innovation Type and Firm Growth Rates

E1 More cited patents lead to higher growth for a firm. This effect is larger for small firms.

E2 An external patent leads to higher growth than an internal patent on average.

E3 More R&D and patent intensive firms grow faster.

E4 Everything else equal, firms that obtain more external patents are more likely to survive. Firms that receive more external citations are more likely to exit the economy.

3 Empirics of Innovation

This section describes empirical regularities regarding innovation and the firm size distribution that come from our model. The presentation follows the five categories outlined above. We touch lightly on the moments that are well established in the literature, devoting relatively more attention to our new contributions. The appendix provides a more extended discussion of these empirics, including those in Figure 1, and appropriate references.

3.1 Data Development

Our project employs the Longitudinal Business Database, the NSF Survey of Industrial Research and Development, and the NBER Patent Database. The Longitudinal Business Database (LBD) provides the backbone for our research. This business registry contains annual observations for every private-sector establishment with payroll from 1976 onward. The Census Bureau data are an unparalleled laboratory for studying the firm size distribution, entry/exit rates, and life cycles of US firms. Sourced from US tax records and Census Bureau surveys, the micro-records document the universe of establishments and firms rather than a stratified random sample or published aggregate tabulations. As a representative year, the data include 108 million workers and 5.8 million establishments in 1997.

The Survey of Industrial Research and Development (RAD) is the US government's primary instrument for surveying the R&D expenditures and innovative efforts of US firms. This is an annual survey conducted jointly by the Census Bureau and NSF. Foster and Grim (2010) provide a comprehensive description of these data. The survey includes public and private firms, as well as foreign-owned firms, undertaking over one million dollars of R&D within the US. RAD surveys are linked to the LBD through Census identifiers. These micro-records begin in 1972 and provide the most detailed statistics available on firm-level R&D efforts. In 1997, 3,741 firms reported positive R&D expenditures that sum to $158 billion.

To complement the RAD, we also match patent data into the LBD. We employ the individual records of all patents granted by the United States Patent and Trademark Office (USPTO) from January 1975 to May 2009. Each patent record provides information about the invention and the inventors submitting the application. Hall et al. (2001) provide extensive details about these data, and Griliches (1990) surveys the use of patents as economic indicators of technology advancement. We only employ patents 1) filed by inventors living in the US at the time of the patent application, and 2) assigned to industrial firms. In 1997, this group comprised about 45 thousand patents (40% of the total USPTO patent count in 1997). We match these patent data to the LBD using firm name and location matching algorithms that build upon Balasubramanian and Sivadasan (2010) and Kerr and Fu (2008). The appendix describes this matching procedure and all the data employed more extensively.

Our empirical approach below is to use the best data and econometric frameworks that we can for each model prediction. We often separately analyze Census Bureau data from external patenting data that do not require disclosure of results. This empirical strategy entails that the samples below differ somewhat across the tests. Nonetheless, this approach delivers the most reliable estimates of each parameter, and we have further confirmed that the reported results are consistent with each other within the samples whenever possible.

3.2 Firm Size Distribution and Firm Growth Rates

Theoretical predictions A1-A5 are well known facts about the firm size distribution that we do not replicate in detail. These regularities closely relate to Stylized Facts 6-10 from Klette and Kortum (2004).[13] They are documented extensively in surveys such as Sutton (1997), Caves (1998), and Geroski (1998). Each of these facts is separately confirmed in the US Census Bureau data that we employ (e.g., Dunne et al. 1988, Davis et al. 1996). They hold across

---

[13]The Klette and Kortum (2004) model displays Gibrat's Law, but they note observed departures from proportionate growth for firms in the stylized facts.

the universe of US firms and across the R&D producing firms that are the focus of this study (e.g., Panel 1B of Figure 1 provides direct evidence on A2-A5). The faster growth for small firms (i.e., departures from Gibrat's Law of proportionate growth) is evident in our firm sample discussed in Table 7 below.

3.3   Firm Size Distribution and Innovation Intensity

Even before separating exploration R&D from exploitation R&D, our model makes specific predictions in B1-B3 about the general extent of R&D and patenting in the firm size distribution. Patents counts and R&D expenditures are predicted to increase with firm size, while their intensities per employee or by firm sales are predicted to decline with firm size. Table 1 documents these patterns using data from 1993 to 1997 for US firms with R&D expenditures. We choose this period because the overlap and quality of our data sources (especially for sales) are strongest during this five-year interval. Expenditures are in constant 1997 dollars. Panels A and B use employment and establishment counts for size categories, respectively.

Columns 1-3 show that, not surprisingly, average R&D expenditures and patent counts grow with firm size. This connects to the first part of B1 and B2. As discussed further in the appendix, the patent bridge files employ name matching procedures that require choosing how conservative or aggressive to be in establishing a link (e.g., how similar two names are before a match is declared). We report results using both our most conservative and most aggressive strategies in Table 1. When an R&D performer is not matched to the patent data, we infer that no patenting occurred for the firm during the period. Both techniques paint a similar picture, and our subsequent empirical work employs the more conservative technique.

Columns 4-6 report the intensity of innovation normalized by employment counts. Column 7 reports intensity of innovation normalized by firm sales. For these intensity patterns, we trim the 1% tails to guard against outliers excessively influencing the results. While larger firms undertake more R&D and obtain more patents, conditional intensities show a strong decline with firm size. This decline is even steeper without the trimming procedure. This connects to the second part of B1 and B2.

These patterns differ somewhat from earlier work that concluded that conditional intensity is independent of size. The appendix undertakes an extended analysis of the firm size to R&D relationship across the 1975-1999 period. When using 1970s data, we find a substantially flatter size relationship that mirrors the invariance to firm size found in the central studies of Cohen (1995) and Cohen and Klepper (1992, 1996a). We show how the pattern has consistently tilted over the past three decades. For example, a cross-sectional regression of log firm R&D on log firm size yields declining coefficients with later cohorts. Elasticities are typically greater than 0.9 in the 1970s and fall to 0.6 in the late 1990s.

This extended discussion also provides additional stylized facts about the complete firm size distribution that are beyond our model of innovative firms. For example, across all firms, the intensity of R&D expenditures or patenting increases with firm size. This is very natural as the calculation includes the vast majority of small firms that do not undertake innovation efforts, for example inferring zero R&D efforts for restaurants, law firms, and computer repair shops. We do not model these firms in our work, typical of endogenous growth frameworks. As the probability that an existing, non-innovative firm commences innovation over the ensuing five years (conditional on survival) is only about 1%, this exclusion seems reasonable.

Our final prediction B3 is that younger firms are more R&D intensive than older firms

among innovative firms. In our model, this is due to the positive relationship between firm age and firm size. This feature is also present in the data, but we do not separately tabulate it since the size relationship is the determining factor.[14]

3.4 Patent Citation Behavior and Innovation Spillover Size

Our model provides the micro-foundations for a simple, powerful test in Lemma 2 as to whether the spillover benefits from exploration R&D are on average stronger or weaker than from exploitation R&D. Figure 2 illustrates this test using patent citations from the NBER Patent Database.[15] The sample is restricted to industrial firms that have all inventors located in the US. We plot the distribution of external citations received over 1975-1999 for patents filed between 1975-1984.

The solid line represents patents that make no backwards citations to prior work of the assignee at the time of filing, a proxy for exploration R&D. The dashed line represents patents whose internal backwards citations are the majority of their citations, a proxy for exploitation R&D. Galasso and Simcoe (2010) similarly use self citation behavior to distinguish exploration-exploitation efforts. Both series display a large number of patents with no external citations (C1) and a skewed distribution (C2).

More important are the comparisons between exploration and exploitation patents, which confirm C3 and provide the answer as to which part of Lemma 2 holds. The spillovers of exploration innovation exceed exploitation innovation akin to first-order stochastic dominance. The differences are statistically significant and hold in regressions that control for a variety of traits about the patents (e.g., technology-year fixed effects) or assignee fixed effects. The omitted, middle group (i.e., patents where backwards self citations are present but a minority) behaves similarly to the no self citation group and are excluded for visual clarity.

Our model identifies a comparative advantage for entrepreneurs and small firms in undertaking exploration innovation due to increased managerial attention of large firms on refining their existing product lines. Under Figure 2's conditions, small firms and new entrants also yield greater spillovers into economic growth due to these forces.

Our model also highlights that the distribution of patent citation life is highly skewed (C4). Figure 3 shows the declining citation rates with patent age. We select three cohorts: granted patents applied for in 1975, 1977, and 1979. Figure 3 graphs the count of citations to these patents in subsequent years normalized by the contemporaneous number of patents. This normalization removes secular changes in patent grant rates. The general patterns in Figure 3 do not reconcile with earlier models of endogenous growth like Romer (1986, 1990). In its most basic form, a "standing on the shoulders of giants" type framework predicts constant citation rates over time for an invention relative to the flow of new inventions.[16,17]

---

[14] Acemoglu et al. (2010) consider innovation dynamics jointly with respect to firm size and age. Foster and Grim (2010) provide related empirical analyses of R&D efforts in the firm size and age distributions.

[15] Hall et al. (2001) provide a comprehensive introduction to patent citations. See also Trajtenberg (1990), Jaffe et al. (2000), and Hall et al. (2005).

[16] The model makes two predictions with respect to age differences across citation types. First, the expected age of internal citations should be less than external citations. This is very stark in the data at 5.9 years versus 8.6 years. Second, and more weakly, the expected average of citations to external patents should be greater than internal patents. Empirically, this difference is small at 8.2 versus 7.8 years.

[17] We follow Mehta et al. (2010) in calculating patent age as the time from the grant date of the cited patent to the application date of the citing patent. If instead using the application date of the cited patent, citation

3.5 Innovation Type and Firm Size Distribution

We next examine variations in innovation type and patent citations across the firm size distribution. We build most directly on the important work of Acs and Audretsch (1987, 1988, 1991), Baumol (2009), Zucker et al. (1998), Kortum and Lerner (2000), Samila and Sorenson (2010), and similar. These papers find a special relationship between small, entrepreneurial companies and major inventions, but they do not study innovation properties systematically throughout the firm size distribution. These distinctions also closely relate to the product versus process innovation differences discussed in the introduction.[18]

We document these relationships in Tables 2 and 3 using estimations at the assignee and firm levels, respectively. For assignee-level estimations, we measure assignee size as the number of granted patents from the USPTO during the 1975-1984 period. For firm-level estimations, we measure firm size as the average number of employees during this period. We employ the earlier period to allow for a long horizon to measure citations.

Each approach has merits and limitations. Patent assignees perhaps most closely resemble the model's direct link of firm size to firm innovation aggregates. This strong link is due to assignees data only considering innovative traits of firms. Using assignees also has the advantage of not requiring Census Bureau clearance of results, which enables the graphs below. On the other hand, many firms file patents through multiple assignees. A firm-level approach better captures these corporate structures. Second, and most important, the direct link of innovation to realized employment growth is of first-order importance.

In Tables 2 and 3, the first panel presents a linear specification of the form,

$$\text{Cite}_p = \eta_{i,t} + \beta \cdot \text{FirmSize}_p + \varepsilon_p,$$

where p indexes patents. The dependent variable is the citation behavior of the patent, which varies across specifications, and the primary regressor is the size of the firm filing the patent. We include $\eta_{i,t}$ fixed effects for the technology i and year t of the patent. Technologies are defined through patent classes. These fixed effects remove systematic differences in citation rates and firm sizes across technologies and years, focusing just on within variations. Estimations are linear probability models, and we find similar results with non-linear techniques.[19]

The second panels in Tables 2 and 3 present non-parametric specifications that include indicator variables for firm size categories. Coefficients in these estimates are relative to the omitted category of the smallest firms. For the assignee estimations, the omitted group is assignees with one patent during the 1975-1984 period. For the firm estimations, the omitted group is firms with fewer than 100 employees on average per annum. The smallest group typically accounts for a large number of firms and a modest share of patents.[20]

---

rates tend to rise for a few years before declining steadily. Patent grant dates typically lag application dates by a few years, and patents become more generally publicized after their grant.

[18] Our work likewise relates to the economics literatures on innovation and industry structure and evolution. Examples include Jovanovic (1982), Nelson and Winter (1982), Gilbert and Newbery (1982), Reinganum (1984), Klepper and Grady (1990), Hopenhayn (1992), Henderson (1993), Jovanovic and MacDonald (1994), Lerner (1997), Gans et al. (2002), Cabral and Mata (2003), Duranton (2007), and Kerr (2010). A lengthy management literature further discusses these ideas (e.g., Christensen 1997).

[19] We weight patents such that each firm or assignee receives equal weight. We also cluster standard errors by firm to account for the multiple patent-to-firm mappings. We find very similar results when collapsing the data to a firm-level cross-section or to a cross-section of firm sizes. Analyzing the data at the patent level, however, allows us to directly control for technology-year fixed effects and undertake additional extensions.

[20] For example, firms with 100 employees or fewer account for 61% of matched firms and 9% of matched

3.5.1   Exploration versus Exploitation Investments

We begin with variations across firm sizes in the types of innovations pursued. Our model links exploitation investments with patents that do not receive external citations. Column 1 of both tables shows that the likelihood of not receiving an external citation rises for patents of larger firms (prediction D1). The relative probability is 1.5% higher for assignees with more than 1000 patents compared to assignees with 101-1000 patents. Likewise, the probability is 1.9% higher for firms with more than 5000 employees compared to firms with 1500-5000 employees. There is also a mostly upward slope in probability across the firm size distribution, although the deviations of the largest firms are most pronounced.

The same relationship holds in Table 4 for the fraction of citations that are self citations of the firm's previous work (D2), which our model also connects to exploitation efforts. This table considers patterns for patents filed in 1995 and their citations over the previous five years. This short period facilitates simulations discussed next, and this snap shot is very representative of the general behavior across the full sample. In 1995, the self citation share grows from 9% for firms filing just one patent to 17% for firms filing 2-5 patents. The share further increases to 31% for firms filing over 100 patents.

The last three columns of Table 4 evaluate these observed self citation shares against counterfactuals. Large patenters are more likely to cite themselves due to the greater likelihood that they draw upon their past work. This is true even if citations are random. If IBM and a small firm in 1995 draw a random citation for the computer industry from 1990-1995, the likelihood that IBM draws itself is much greater. The likelihood of self citing for a new entrant is naturally zero. This bias to firm size is particularly true where large firms dominate narrow technology fields.

To confirm that this mechanical effect is not driving the observed relationship in Column 2, we undertake Monte Carlo simulations where we replace observed patents with random counterfactuals. For each observed citation, we draw a counterfactual that matches the technology and application year of the cited patent. We include the original citation among the possible pool of patents, and we draw with replacement. We measure from the simulation a counterfactual self citation share to assignee size relationship. As this relationship depends upon the randomness of the simulation draws, we repeat the procedure 1000 times.

We use these 1000 simulations to generate 95% confidence bands for the self citation ratio of each assignee. These confidence bands are specific to assignees based upon their size and underlying technologies. These confidence bands more rigorously test whether the observed self citation relationships are a systematic departure from the null hypothesis of being randomly determined. As anticipated, Column 3 shows that the mean value of the test statistic is rising in firm size.

Columns 4 and 5 confirm that the observed self citation behavior is a significant departure among large assignees. Column 4 examines the prevalence of departures. For assignees with one patent during 1995, only 13% display self citation behavior that we can reject as being random at a 95% confidence level. This non-random share grows to 97% for assignees with more than 100 patents in 1995. Column 5 also shows that average deviation of self citation shares from the random baseline is growing in firm size. These departures indicate that our

---

patents. The largest firm size category accounts for 4% of matched firms and 72% of patents. We discuss in the appendix how our conservative matching approach results in these shares being a lower bound on the small firm contribution. Likewise, over 50% of assignees have one patent, and 90% have fewer than 10 patents.

results are due to firm behavior rather than the mechanics of firm size. The appendix further shows that these self citation findings hold in within-firm panel analyses, too.[21]

3.5.2 Qualities of Exploration Innovations

These first two regularities (D1, D2) describe differences by firm size in whether exploration or exploitation innovations are pursued. Before proceeding to predictions D3-D6, which relate to the complete distribution of patent qualities, we pause to consider whether exploration innovations differ in quality for small and large firms. Assumption (3) specifies a common research technology for exploration innovations. This suggests that the citation qualities of exploration inventions should not vary significantly with firm size (exact predictions depend upon model parameters).

Figures 4 and 5 provide simple evidence using exploration patents that receive an external citation. We plot the mean, non-zero external citations across exploration patents for assignees against their sizes. Figure 4 uses a non-logarithmic citation scale, while Figure 5's scale is logarithmic. The trend lines employ lowess functional smoothing. In both cases, we find broad similarity in patent quality and citations across the firm size distribution for external advances. This verifies assumption (3). Differences in types of innovation pursued across the firm size distribution are more important than differences in research technologies.

3.5.3 Complete Quality Distributions

To combine these effects and study the complete distribution of innovations, we categorize patents by quartiles of external citations received. Our model identifies how external patent citations reflect the underlying step sizes or qualities of innovations. Thus, our analysis of the patent citation distribution also describes the distribution of innovation qualities.

We develop these quartiles by technology-year using the 36 sub-categories of the USPTO classification scheme (e.g., "Optics", "Biotechnology"). Quartiles are assigned by patent, and patents with no external citations are assigned zero values for identifying bands. Tables 2 and 3 continue with our 1975-1984 sample of industrial patents. The dependent variables in Columns 2-5 are indicator variables for quartiles of the citation distribution. The lowest quality quartile (0%-24%) is Column 2, and the highest quartile (75%-100%) is Column 5. Coefficients across Columns 2-5 naturally sum to one.

Both tables show that the patents of larger firms are more likely to be of lower citation quality. This is primarily a consequence of shifting mass from the upper two quartiles of the distribution to the lowest quartile, and in particular through patents with no external citations. These specifications thus confirm our earlier observations in a flexible way that connects with our firm growth estimations shortly. They directly connect to predictions D3 and D4.

We can also use this framework to further confirm our hypothesis that innovation quality is mostly stochastic conditional on innovation type. Prediction D4 suggests that the innovations of small firms are more likely to be of higher quality, but this differential should continually weaken as more stringent citation quality thresholds are examined as the stochastic nature of

---

[21] This analysis closely relates to the patent localization work of Jaffe et al. (1993) and Thompson and Fox-Kean (2005). Similar procedures are used in agglomeration calculations like Duranton and Overman (2005) and Ellison et al. (2010). Agrawal et al. (2010) discuss related issues with respect to large patenting firms in "company towns" and their self citation behavior (e.g., Eastman Kodak in Rochester, NY).

realized inventions becomes more important. These weakening differences through the upper tail of the quality distribution are a unique prediction of Proposition 7 that helps differentiate our model from other frameworks.[22]

To examine this second prediction, we look within the upper quartile of the citation distribution in Columns 6-8 of Tables 2 and 3. We continue with indicator variables for whether a patent is above a certain threshold in the quality distribution. Unlike Columns 2-5, however, we normalize the dependent variable to have unit standard deviation so that coefficients are comparable across Columns 6-8. This step was not required earlier as the quartiles are approximately of equal size.

When using assignee data in Table 2, the lower arrival rate of major innovations for the largest firms compared to the smallest firms persists to the 99th percentile of patent quality. The firm size differential does, however, lose about two-thirds of the strength that existed when looking at the top quartile. When using firm data in Table 3, the small firm advantage is more prominent in the third quartile. What advantages exist at the 75th percentile are entirely eroded by the 95th percentile. These patterns support the more subtle part of prediction D4.

The appendix further shows these results hold when examining patent claims. Each patent includes a series of claims that delineate the property rights and novel features of the technology. While we develop our predictions in Section 2.8 for patent citations, most distributional results also hold for patent claims as claims are connected with the underlying step sizes of innovations. Claims are also linked by prior empirical work to the quality of innovations, and so from an empirical perspective it is reassuring that our results hold under this alternative technique.

We conclude that small firms are more likely to undertake major innovations. This is primarily due to differences in types of innovations pursued. When undertaking exploration R&D, quality differences by firm sizes are very limited, especially at the upper tail of the quality distribution where the stochastic nature of invention is critical in shaping outcomes.

3.5.4   Dynamic Effects

Under the conditions of Lemma 2, predictions D5 and D6 emerge about citations profiles of patents by firm age. We test these predictions using industrial patents from 1977-1994. This sample gives us sufficient history in the patent and Census Bureau data to be confident in first observations being new entrants. Ending the sample in 1994 allows for some time to measure subsequent citations. We observe D5 directly, for example, by calculating age for assignees through the time since their first patent filing. Patents for new entrants receive about 25% more external citations than patents for firms observed for ten years (6.8 versus 5.4). The declines in citations are monotonic in between.

For D6, Table 5 also presents some simple panel evidence on patent quality within firms over time. We restrict the sample to new entrants during 1977-1994. We regress traits of patents on an indicator variable for whether or not the patent is filed in the first two years that a firm is observed in the patent data. We include assignee or firm fixed effects to compare

---

[22] A simple example can illustrate this point. Consider a scenario where 100 patents for small and large firms are drawn from a uniform quality distribution [0,1]. Imagine too that we include an additional 100 large firm patents that are of zero quality. Two consequences are 1) the mean of the small firm distribution will usually be higher than the mean of the large firm distribution, and 2) there is no greater likelihood that the best patent is made by a small company. Because the randomness of the invention draws determines the extreme values realized, a similar logic applies to the very highest thresholds of patent quality.

early patents of the firm to later patents. We also include technology-year fixed effects.

Column 1 shows that the average external citation count is higher at entry. Column 2 shows that patents also have larger numbers of claims at firm entry than in later years. Columns 3-6 show the distribution of external citations in quartiles. Column 3 is the lowest quality quartile, and Column 6 is the highest quality quartile. Entrants have disproportionate representation in the highest quality quartile compared to later years for the same firm. The results confirm the time path of firms in terms of invention quality.

Our remaining dynamic evidence focuses on verifying our model's assumption that major exploration innovations are followed within firms by more incremental innovations and refinements. This process requires that an external innovation be made to dramatically push forward the technology of a product line that is dominated by incremental inventions within the currently leading firm. We can further verify these features by demonstrating that the mean quality of citing patents outside of the original firm for a given invention is higher than the mean quality of citing patents within the firm.

Table 6 describes differences in how internal versus external work build upon an invention. We use a linear specification of the form,

$$\text{Cite}_{p_2,p_1} = \phi_{p_1} + \eta^{p_2}_{i,t} + \beta \cdot \text{External}_{p_2,p_1} + \epsilon_{p_2,p_1},$$

where $\text{Cite}_{p2,p1}$ models traits of patents p_2 that cite patents p_1. We include citations for US industrial patents filed during 1975-1984. We restrict the citations to be US industrial patents filed within a ten-year window of the original patent. We find similar patterns when using all citations, but the consistent window is more appropriate.

The primary regressor is the indicator variable $\text{External}_{p_2,p_1}$ that takes unit value if the assignee of citing patent p_2 differs from the assignee of cited patent p_1. Three-quarters of citations are external. We include ϕ_{p_1} fixed effects for cited patents. We thus compare differences between internal and external citations on the same patent. We also include $\eta^{p_2}_{i,t}$ fixed effects for the technology i and year t of the citing patent p_2; the patent fixed effects naturally control for these traits for cited patents p_1. We define $\eta^{p_2}_{i,t}$ through USPTO sub-categories and five-year time periods. We cluster standard errors by cited patents.

The first column of Table 6 models the number of external citations on citing patents p_2 as the outcome variable. The second column alternatively tests the number of claims on the citing patent as a measure of quality. Columns 3-6 then test the quality distribution of citing patents in a format similar to Tables 2 and 3. Quality distributions are determined through ranks of external citations by technology and period. Coefficients across the final four columns for a row approximately sum to zero, but the relationship does not hold exactly given that quality distributions are calculated over a larger group than the regression sample.

The first column finds that the mean number of future citations for external work that builds upon a given invention is 0.8 citations higher than the internal work that also builds on the focal invention. This effect is large relative to the sample mean of 8.2. There is also a substantial external premium of 1.2 claims relative to the sample mean of 15.4. Columns 3-6 show that this effect mainly comes from a greater prevalence of upper quartile patents among the external citing patents, with mass moved from the lowest two quartiles of the distribution. These patterns suggest that external work that builds upon a given invention is stronger than the internal work that follows. The appendix extends this analysis.[23]

---

[23] The appendix also reports transition matrices for firms in the patent quality and patent count distributions.

3.6   Innovation Type and Firm Growth Rates

Our final group of predictions E1-E4 focuses on the link between innovation and firm growth. We show these relationships using the full dataset of patents matched to the Census Bureau data. We organize our sample around five-year blocks. The three periods included in the regressions are 1978-1982, 1983-1987, and 1988-1992. We calculate for each firm its employment growth to the following period, and we use 1993-1997 data to calculate growth for 1988-1992.

Following Davis et al. (1996), we calculate employment growth relative to the average of the two periods: $EmpGr_{f,t} = [Emp_{f,t+1} - Emp_{f,t}]/([Emp_{f,t+1} + Emp_{f,t}]/2)$. This growth variable is bounded between (-2, 2) for continuing firms, reduces the impact of outliers, and protects against mean reversion. For our core sample, the mean of $EmpGr_{f,t}$ is 0.128 and the standard deviation is 0.597. Firm employment $Emp_{f,t}$ for each period is calculated as a mean across the years where positive employment is observed. Our patterns below are robust to several variants of defining employment growth (e.g., $\ln[Emp_{f,t+1}/Emp_{f,t}]$).

The central regressors to explain employment growth to the next period are the firm's current employment, the firm's total patenting in the period, the quality distribution of the firm's own patents in this period (Internal Patent Quality Share$_{f,q}$), and the quality distribution of the recent external innovations cited by the firm (External Citations Quality Share$_{f,q}$). External citation quality is calculated with a window of five years.[24]

Table 7 presents an employment growth specification of the form,

$$
\begin{aligned}
EmpGr_{f,t} = {} & \eta_{i,t} + \gamma_E \ln(\mathrm{Emp}_{f,t}) + \gamma_P \ln(\mathrm{Patents}_{f,t}) \\
& + \sum_{q \in Q} (\beta_q \cdot \text{Internal Patent Quality Share}_{f,q}) \\
& + \sum_{q \in Q} (\theta_q \cdot \text{External Citations Quality Share}_{f,q}) + \epsilon_{f,t},
\end{aligned}
$$

where f and t index firms and five-year periods. The set of quality quartiles Q are indexed by q. We omit the lowest two quality quartiles for both distributions, and the β_q and θ_q coefficients are estimated relative to this group. We include $\eta_{i,t}$ fixed effects for the industry i and year t of the firm. Industries are assigned to firms at the two-digit level of the Standard Industrial Classification system using industries in which firms employ the most workers. We cluster standard errors at the firm level; regressions are unweighted.[25]

Similar to our model, these matrices emphasize the faster speed of firm movements across the citation quality distribution compared to the patent count distribution.

[24] For a firm observation during the 1983-1987 period, we first calculate the fraction of its 1983-1987 patents that fall within each of the four quality quartiles. For the firm's patents filed in 1987, we also separately calculate the quality distribution of cited, external patents during 1982-1987. We repeat this procedure for patents filed by the firm in 1986 using a 1981-1986 window on external citations, and so on. We then aggregate these external citations qualities to characterize for the firm the overall quality of the prior external work upon which it was building during the 1983-1987 period. For the first sample period of 1978-1982, we use data on patents from before 1978 to calculate quality. The five-year window allows for consistent calculations throughout the panel. Quality distributions are relative to the technology-year in which a patent was filed.

[25] This growth specification implicitly contains firm fixed effects. First, the dependent variable is employment growth versus the level of employment. Second, the innovation literature often characterizes patents as stock variables. By focusing on contemporaneous patent counts and quality distributions, we are implicitly looking at changes in innovation stocks to the current period from the previous period. Econometric specifications tend to find contemporaneous R&D investments have the most important impact for rates of technology formation

The first column of Table 7 shows the employment growth response when controlling for just employment and patenting levels. The negative relationship to current employment levels matches the departures from Gibrat's Law recently observed in complete datasets of the firm size distribution (prediction A3). If Gibrat's Law were to hold throughout the full size distribution, the γ_E employment coefficient would be zero. Instead, small firms grow faster. The estimate suggests that a 10% increase in firm size lowers employment growth by 5% from the sample mean of 0.128. Column 1 also finds that greater numbers of patents promote growth in the next period. Non-linear specifications find the employment relationship is convex, but adding higher order moments does not influence the quality distribution results reported next.

Column 2 introduces the quality distributions: Internal Patent Quality Share$_{f,q}$ and External Citations Quality Share$_{f,q}$. The β and θ coefficients are measured relative to the two lowest quality quartiles. Looking first at own-firm patents, employment growth is higher when a greater share of the firm's own innovations are higher quality. The effect is statistically significant and economically important. Moving 10% of the quality distribution from the lowest quartile to the highest quartile has the same employment growth effect as a 12% increase in the number of patents. This both supports our model's assumptions and prediction E1.

The spillover effect captured by citations of high quality patents is also strong for the upper quality distribution. Column 2's θ coefficient for the upper quality quartile is greater than its corresponding β coefficient, although in general we find that the θ and β coefficients for the upper quartile are comparable in size. This suggests that building upon high quality innovations is very important for growth. Independent of magnitude, we also consistently observe more nonlinear differences in the quality distribution for external citations compared to internal patents (i.e., comparing the third and highest quality quartiles). Spillover benefits are highly skewed towards high quality innovations.[26]

The next two columns split the sample into small firms and large firms. If a firm is observed during the sample period to have 100 or fewer employees, it is classified as a small firm for the full sample. Not surprisingly, deviations from Gibrat's Law are greater for small firms. Growth is also stronger in total patent counts for small firms compared to large firms. Finally, the growth impact of developing a major innovation or of building upon a prior major innovation is more powerful in small firms (prediction E1).

Our primary sample drops firm-year observations where a patent is not observed due to the log transformation. This truncation cannot be remedied by selection models or examining raw patent counts because the associated quality distributions are undefined. Column 5 takes the alternative approach of restricting the sample to firms that are observed to patent in all three periods, circumventing the selection bias within firms. The overall importance of having major innovations within the company grows relative to the quality of external citations in this sample. In both cases, high quality innovations remain particularly important.

The last column separately confirms prediction E2 that internal patents do not contribute

---

(e.g., Pakes and Griliches 1980, Hausman et al. 1984, and Hall et al. 1986). Samila and Sorenson (2010) likewise find localized employment effects have very short lags.

[26]These results, of course, should not be interpreted as suggesting an easy way to promote firm growth. Adjusting external citation quality distributions is a very difficult challenge for managers, perhaps even more so than internal quality distributions. Breakthrough innovations are often difficult to discern, and how to build upon them appropriately is not obvious. In this context, it is important to note that USPTO examiners review and sometimes modify citations on patents to reflect appropriate prior technologies. Citations thus reflect true technological proximity versus an attempt to actively define how a firm is being viewed externally.

as much to firm growth as external patents. We calculate the share of citations made by the firm in the current period that are self citations of prior work by the firm. We include indicator variables for the self citation share being (0%, 20%] and (20%, 100%]. Relative to the majority of observations that have zero self citations, higher backwards self citations are associated with weaker employment growth. This relationship holds in all subsamples and when only including the self citation measures.

Prediction E3 also holds with this sample. Regressing employment growth on log patenting by itself yields a coefficient of -0.018 (0.003). This negative relationship is to be expected given that greater patenting is associated with larger firm size. On the other hand, regressing employment growth on log patent intensity by itself yields a coefficient of 0.066 (0.002). Similar relationships hold for R&D expenditures using the samples in Table 1.

The appendix reports several specification extensions. We again repeat our analysis with patent claims as a second measure of innovation quality. We also test the stronger employment growth effects associated with higher patent quality threshold effects. Finally, we show that survival margins confirm prediction E4.

4   Conclusions

Research continues to enrich our models of economic growth. This paper contributes to this effort by modelling heterogeneity in firm behavior for exploration versus exploitation R&D. We find empirical evidence that exploration R&D does not scale as rapidly with firm size as exploitation R&D. Adding this ingredient to our theoretical framework delivers a richer account of innovation and growth, offering many testable implications that hold in the data. The R&D heterogeneity across different-sized firms is important.

One important implication of the model is that small firms, and especially new entrants, have a comparative advantage for undertaking exploration R&D. This is very natural as large firms have many product lines to concentrate on. This framework rationalizes why small, entrepreneurial firms contribute disproportionate numbers of major innovations. Moreover, the predicted distributional patterns line up with the data.

Another important implication emerges from patent citations. By incorporating patent citations behavior into our model, we derive a simple test as to whether the growth spillover effects across firms are higher for exploration or exploitation innovation. For the recent US economy, we find that the external impacts of exploration innovation on other firms exceed exploitation innovation. This in turn suggests that small, innovative firms and new entrants play a special role in economic growth due to these spillovers.

This paper establishes a theoretical framework to study innovative firms and provides empirical evidence from US firms. In a parallel quantitative work, we are calibrating a version of this model to US data that also incorporates non-innovating firms at the expense of losing theoretical tractability. That work will quantify the implications of different R&D types on firm dynamics, and, more important, the contribution of innovating firms to the overall economy. Our ultimate goal in that project is to study quantitatively various innovation policies.

This paper has built a very tractable model with important firm heterogeneities. We hope this framework is a useful platform for future research on the financing of innovation, intellectual property rights, and R&D policies in general. We also hope the new facts generated in our empirical work stimulate future discussions on the dynamics of innovative firms.

References

[1] Acemoglu, Daron, *Introduction to Modern Economic Growth* (Princeton, NJ: Princeton University Press, 2009).

[2] Acemoglu, Daron, and Ufuk Akcigit, "State Dependent Intellectual Property Rights Policy", Working Paper (2010).

[3] Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William Kerr, "Innovation, Reallocation, and Growth", Working Paper (2010).

[4] Acemoglu, Daron, and Dan Cao, "Innovation from Entrants and Incumbents", Working Paper (2010).

[5] Acs, Zoltan, and David Audretsch, "Innovation, Market Structure, and Firm Size", *Review of Economics and Statistics* 69:4 (1987), 567-574.

[6] Acs, Zoltan, and David Audretsch, "Innovation in Large and Small Firms: An Empirical Analysis", *American Economic Review* 78:4 (1988), 678-690.

[7] Acs, Zoltan, and David Audretsch, "Innovation and Size at the Firm Level", *Southern Economic Journal* 57:3 (1991), 739-744.

[8] Adams, James, and Susan Peck, "A Guide to R&D Data at the Center for Economic Studies U.S. Bureau of the Census", CES Working Paper 94-9 (1994).

[9] Aghion, Philippe, Christopher Harris, and John Vickers, "Competition and Growth with Step-by-Step Innovation: An Example", *European Economic Review* 41 (1997), 771-782.

[10] Aghion, Philippe, Christopher Harris, Peter Howitt, and John Vickers "Competition, Imitation and Growth with Step-by-Step Innovation", *Review of Economic Studies* 68:3 (2001), 467-492.

[11] Aghion, Philippe, and Peter Howitt, "A Model of Growth Through Creative Destruction", *Econometrica* 110 (1992), 323-351.

[12] Agrawal, Ajay, Iain Cockburn, and Carlos Rosell, "Not Invented Here? Innovation in Company Towns", *Journal of Urban Economics* 67:1 (2010), 78-89.

[13] Akcigit, Ufuk, "Firm Size, Innovation Dynamics and Growth", Working Paper (2010).

[14] Balasubramanian, Natarajan, and Jagadeesh Sivadasan, "What Happens When Firms Patent? New Evidence from US Economic Census Data", *Review of Economics and Statistics* (2010), forthcoming.

[15] Barro, Robert, and Xavier Sala-i-Martin, *Economic Growth* (Cambridge: MIT Press, 1995).

[16] Baumol, William, "Entrepreneurship, Innovation and Growth: The David-Golliath Symbiosis", Working Paper (2009).

[17] Bernard, Andrew, Stephen Redding, and Peter Schott, "Multi-Product Firms and Product Switching", *American Economic Review* 100:1 (2010), 70-97.

[18] Caballero, Ricardo, and Adam Jaffe, "How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," in *NBER Macroeconomics Annual 1993* (Cambridge: MIT Press, 1993), 15-86.

[19] Cabral, Luis, and Jose Mata, "On the Evolution of Firm Size Distribution: Facts and Theory", *American Economic Review* 93 (2003), 1075-1090.

[20] Cai, Jie, "Knowledge Spillovers and Firm Size Heterogeneity", Working Paper (2010).

[21] Caves, Richard, "Industrial Organization and New Findings on the Turnover and Mobility of Firms", *Journal of Economic Literature* 36 (1998), 1947-1982.

[22] Christensen, Clay, *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail* (Boston, MA: Harvard Business School Press, 1997).

[23] Cohen, Wesley, "Empirical Studies of Innovative Activity", in Paul Stoneman (ed.) *Handbook of the Economics of Innovations and Technological Change* (Oxford, UK: Blackwell, 1995).

[24] Cohen, Wesley, and Steven Klepper, "The Anatomy of Industry R&D Intensity Distributions", *American Economic Review* 82 (1992), 773-799.

[25] Cohen, Wesley, and Steven Klepper, "A Reprise of Size and R & D", *Economic Journal* 106:437 (1996a), 925-951.

[26] Cohen, Wesley, and Steven Klepper, "Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D", *Review of Economics and Statistics* (1996b), 232-243.

[27] Davis, Steven, John Haltiwanger, and Scott Schuh, *Job Creation and Destruction* (Cambridge, MA: MIT Press, 1996).

[28] Dunne, Timothy, Mark Roberts, and Larry Samuelson, "Patterns of Firm Exit and Entry in U.S. Manufacturing Industries", *Rand Journal of Economics* 19 (1988), 495-515.

[29] Duranton, Gilles, "Urban Evolutions: The Fast, the Slow, and the Still", *The American Economic Review* 97:1 (2007), 197-221.

[30] Duranton, Gilles, and Henry Overman, "Testing for Localization Using Micro-Geographic Data", *Review of Economic Studies* 72 (2005), 1077-1106.

[31] Ellison, Glenn, Edward Glaeser, and William Kerr, "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns", *American Economic Review* 100 (2010), 1195-1213.

[32] Foster, Lucia, and Cheryl Grim, "Characteristics of the Top R&D Performing Firms in the U.S.: Evidence from the Survey of Industrial R&D", CES Working paper 10-33 (2010).

[33] Gans, Joshua, David Hsu, and Scott Stern, "When Does Start-up Innovation Spur the Gale of Creative Destruction?", *RAND Journal of Economics* 33 (2002), 571-586.

[34] Geroski, Paul, "An Applied Econometrician's View of Large Company Performance", *Review of Industrial Organization* 13 (1998), 271-294.

[35] Gilbert, Richard, and David Newbery, "Preemptive Patenting and the Persistence of Monopoly", *American Economic Review* 72 (1982), 514-526.

[36] Galasso, Alberto, and Timothy Simcoe, "CEO Overconfidence and Innovation", Working Paper (2010).

[37] Griliches, Zvi, "Patent Statistics as Economic Indicators: A Survey", *Journal of Economic Literature* 28:4 (1990), 1661-1707.

[38] Griliches, Zvi, "The Search for R&D Spillovers", *Scandinavian Journal of Economics* 94 (1992), S29-S47.

[39] Grossman, Gene, and Elhanan Helpman, *Innovation and Growth in the Global Economy* (Cambridge, MA: MIT Press, 1991).

[40] Hall, Bronwyn, Zvi Griliches, and Jerry Hausman, "Patents and R and D: Is There a Lag?", *International Economic Review* 27 (1986), 265-283.

[41] Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools", NBER Working Paper 8498 (2001).

[42] Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "Market Value and Patent Citations", *Rand Journal of Economics* 36 (2005), 16-38.

[43] Hausman, Jerry, Bronwyn Hall, and Zvi Griliches, "Econometric Models for Count Data with an Application to the Patents-R&D Relationship", *Econometrica* 52 (1984), 909-938.

[44] Henderson, Rebecca, "Underinvestment and Incompetence as Responses to Radical Innovation: Evidence from the Photolithographic Alignment Equipment Industry", *Rand Journal of Economics* 24 (1993), 248-270.

[45] Hopenhayn, Hugo, "Entry, Exit, and Firm Dynamics in Long Run Equilibrium", *Econometrica* 60:5 (1992), 1127-1150.

[46] Hopenhayn, Hugo, Gerard Llobet, and Matthew Mitchell, "Rewarding Sequential Innovators: Patents, Prizes, and Buyouts", *Journal of Political Economy* 115:6 (2006), 1041-1068.

[47] Jaffe, Adam, Manuel Trajtenberg, and Michael Fogarty, "Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors", *American Economic Review Papers and Proceedings* 90 (2000), 215-218.

[48] Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *The Quarterly Journal of Economics* 108:3 (1993), 577-598.

[49] Jarmin, Ron, and Javier Miranda, "The Longitudinal Business Database", CES Working Paper (2002).

[50] Jones, Charles, "R&D-Based Models of Economic Growth", *Journal of Political Economy* 103 (1995), 759-784.

[51] Jovanovic, Boyan, "Selection and the Evolution of Industry", *Econometrica* 50 (1982), 649-670.

[52] Jovanovic, Boyan, and Glenn McDonald, "The Life Cycle of a Competitive Industry", *Journal of Political Economy* 102 (1994), 322-347.

[53] Kerr, William, "Breakthrough Inventions and Migrating Clusters of Innovation", *Journal of Urban Economics* 67:1 (2010), 46-60.

[54] Kerr, William, and Shihe Fu, "The Survey of Industrial R&D—Patent Database Link Project", *Journal of Technology Transfer* 33:2 (2008), 173-186.

[55] Kerr, William, Josh Lerner, and Antoinette Schoar, "The Consequences of Entrepreneurial Finance: A Regression Discontinuity Analysis", NBER Working Paper 15831 (2010).

[56] Klepper, Steven, "Entry, Exit, Growth, and Innovation over the Product Life Cycle", *American Economic Review* 86:3 (1996), 562-583.

[57] Klepper, Steven, and Elizabeth Grady, "The Evolution of New Industries and the Determinants of Market Structure", *Rand Journal of Economics* 21 (1990), 27-44.

[58] Klette, Tor Jakob, and Samuel Kortum, "Innovating Firms and Aggregate Innovation", *Journal of Political Economy* 112:5 (2004), 986-1018.

[59] Kortum, Samuel, "Research, Patenting, and Technological Change", *Econometrica* 65 (1997), 1389-1419.

[60] Kortum, Samuel, and Josh Lerner, "Assessing the Contribution of Venture Capital to Innovation", *RAND Journal of Economics* 31:4 (2000), 674-692.

[61] Lentz, Rasmus, and Dale Mortensen, "An Empirical Model of Growth through Product Innovation", *Econometrica* 76:6 (2008), 1317-1373.

[62] Lerner, Josh, "The Importance of Patent Scope: An Empirical Analysis", *RAND Journal of Economics* 25 (1994), 319-333.

[63] Lerner, Josh, "An Empirical Exploration of a Technology Race", *Rand Journal of Economics* 28 (1997), 228-247.

[64] Mehta, Aditi, Marc Rysman, and Tim Simcoe, "Identifying the Age Profile of Patent Citations: New Estimates of Knowledge Diffusion", *Journal of Applied Econometrics*, forthcoming.

[65] Nelson, Richard, and Sidney Winter, *An Evolutionary Theory of Economic Change* (Cambridge, MA: Harvard University Press, 1982).

[66] Pakes, Ariel, and Zvi Griliches "Patents and R&D at the Firm Level: A First Look", *Economics Letters* 5 (1980), 377-381.

[67] Reinganum, Jennifer, "Uncertain Innovation and the Persistence of Monopoly", *American Economic Review* 73 (1984), 741-748.

[68] Romer, Paul, "Increasing Returns and Long-Run Growth", *Journal of Political Economy* 94:5 (1986), 1002-1037.

[69] Romer, Paul, "Endogenous Technological Change", *Journal of Political Economy* 98:5 (1990), S71-S102.

[70] Samila, Sampsa, and Olav Sorenson, "Venture Capital, Entrepreneurship and Economic Growth", *Review of Economics and Statistics* (2010), forthcoming.

[71] Spence, Michael, "Cost Reduction, Competition, and Industry Performance", *Econometrica* 52:1 (1984), 101-121.

[72] Sutton, John, "Gibrat's Legacy", *Journal of Economic Literature* 35 (1997), 40-59.

[73] Thompson, Peter, and Melanie Fox-Kean, "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment", *American Economic Review* 95:1 (2005), 450-460.

[74] Trajtenberg, Manuel, "A Penny for Your Quotes: Patent Citations and the Value of Innovations", *RAND Journal of Economics* 21 (1990), 172-187.

[75] Zucker, Lynne, Michael Darby, and Marilynn Brewer, "Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises", *American Economic Review* 88 (1998), 290-306.

Fig. 1: Exploitation R&D behavior among US firms
Distributions of firm size and growth

1A: Share of firms that self cite past work by firm

Gray: Exiting firms, major declines
Black: Stable employment
White: Entrants, major growth

1C: Share of firms undertaking process oriented R&D

1B: Average share of citations that are self citations

Percentage of observations

1D: Average share of R&D that is process oriented
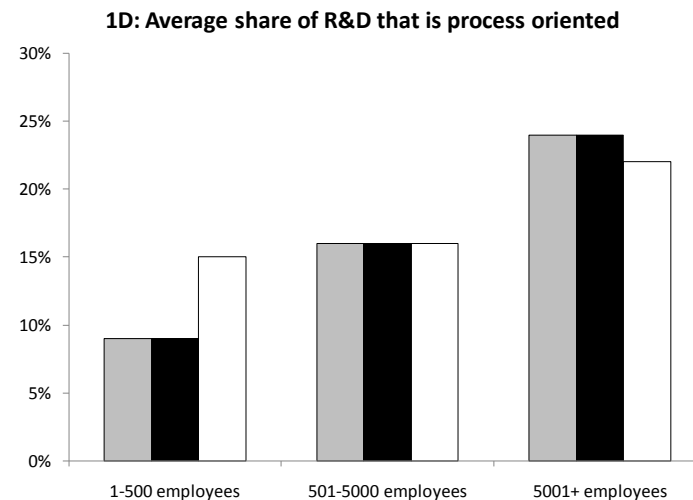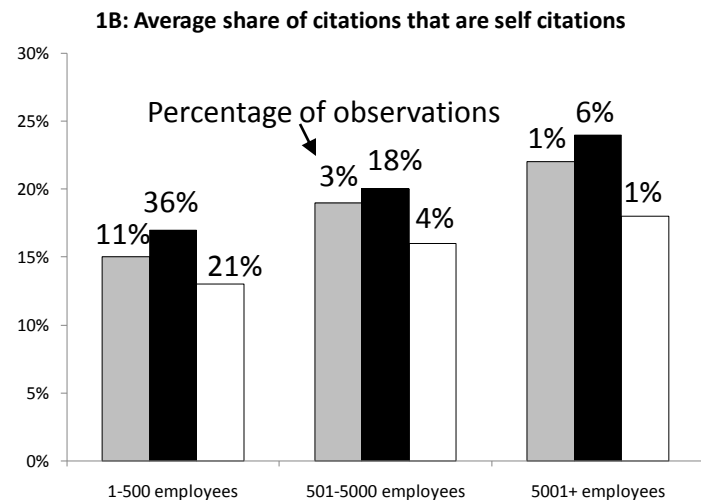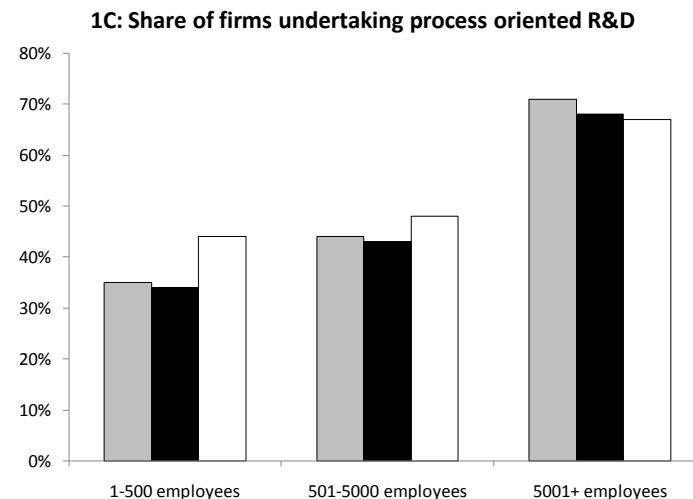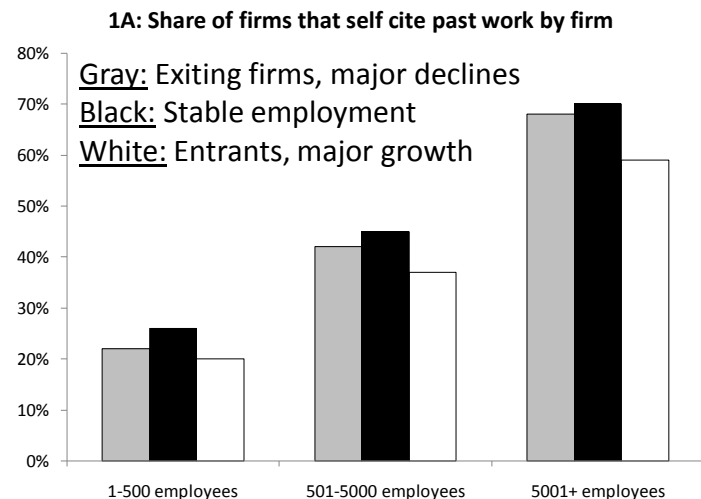
Notes: Figure shows basic regularities on firm R&D and patenting for innovative firms that conduct R&D or file patents. Data are taken from US Census Bureau and NBER Patent Database. The three groups of columns per chart separate firms by employment size. Within each triplet, firms are further separated by contemporaneous employment growth. Panels 1A and 1B document the extent to which firms self cite their prior work. Panels 1C and 1D consider the extent to which firms undertake process-oriented R&D. Greater self citations or process-oriented R&D represent exploitation R&D behavior. Exploitation is increasing in firm size, but is approximately invariant to contemporaneous firm growth rates conditional on firm size.

Fig. 2: Spillovers from external versus internal patents
1975-1984 US domestic patents
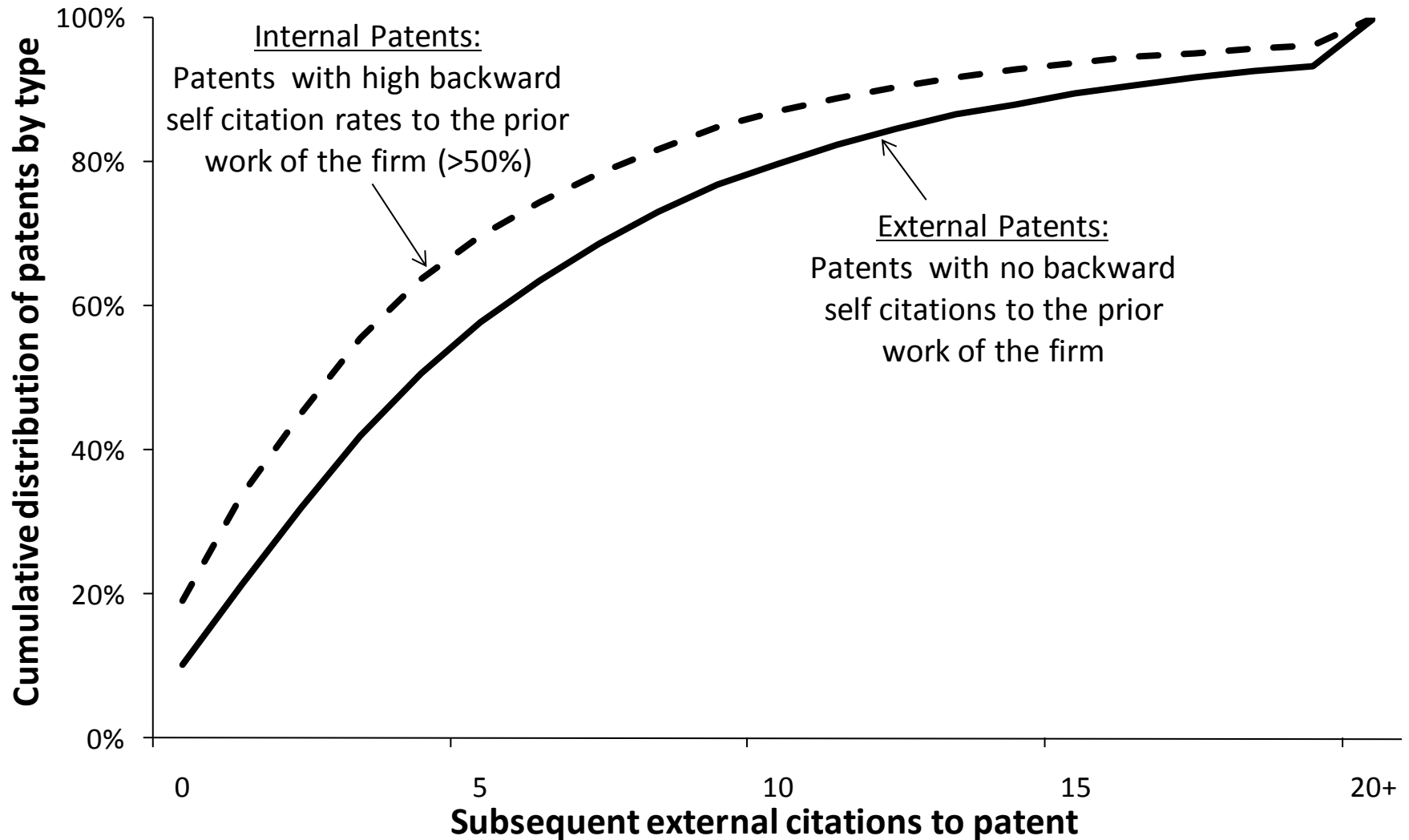
Notes: Figure plots cumulative distribution of citation counts by patent type. We group patents by the share of citations that they make at the time of the patent to prior work by the same assignee. The distribution of citations overall is highly skewed. Patents building mostly on prior work of the same firm have a lower external impact. This is evident in that the cumulative density for external patents is lower than for internal patents. External patents represent exploration innovation, while internal patents represent exploitation. Our sample includes all domestic USPTO patents to industrial firms from 1975-1985. Citations are calculated over 1975-1999.

Fig. 3: Patent citation age lengths
1975, 1977, and 1979 US domestic patents
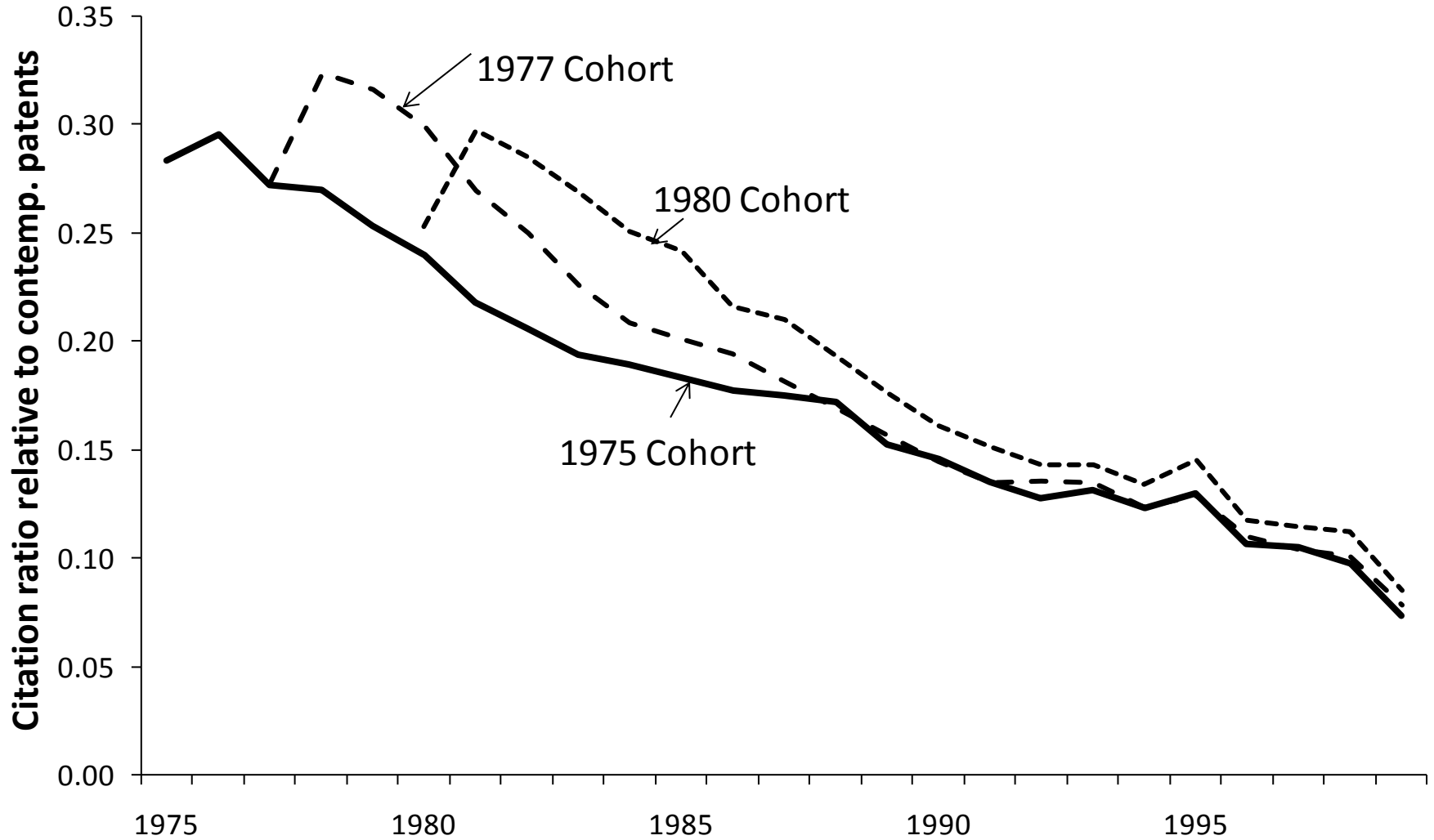
Notes: Figure shows the declining citation rates with patent age. We select three cohorts: granted patents in 1975, 1977, and 1980. We graph the count of citations to these patents in subsequent years normalized by the number of patents in each year. This normalization removes secular changes in patent grant rates.

Fig. 4: External citations and assignee size
1975-1984 US domestic patents
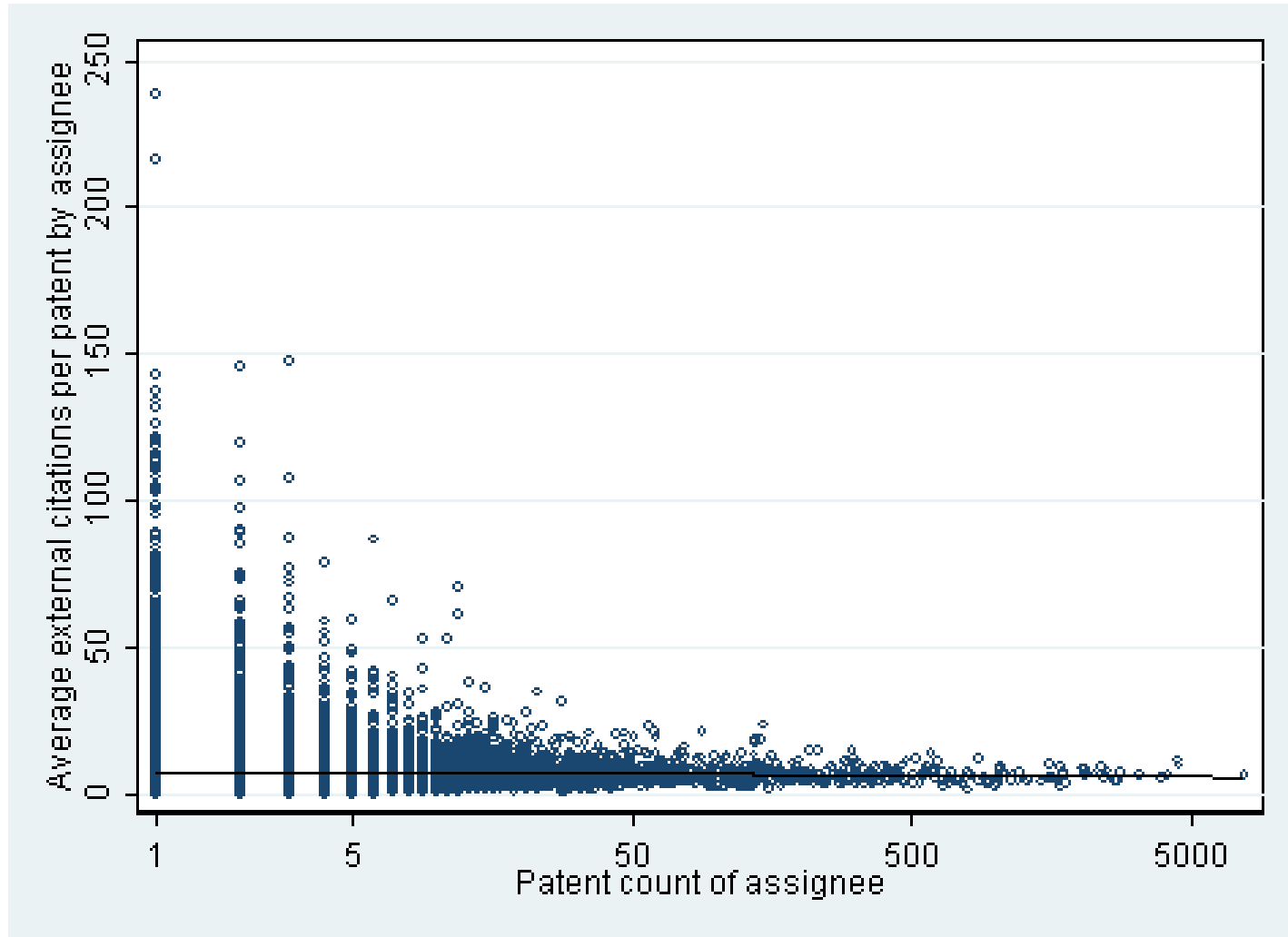
Notes: Figure graphs the average number of external citations per patent against the number of patents for the assignee. Each circle represents an assignee, and the displayed trend line employs lowess functional smoothing. Average external citations are only calculated over patents with positive external citations. The sample includes all industrial patents applied for during the 1975-1984 period. Citations are calculated over 1975-1999. Figure shows how the average number of external citations per patent does not depend strongly on assignee size conditional on being an external innovation.

Fig. 5: Log external citations and assignee size
1975-1984 US domestic patents
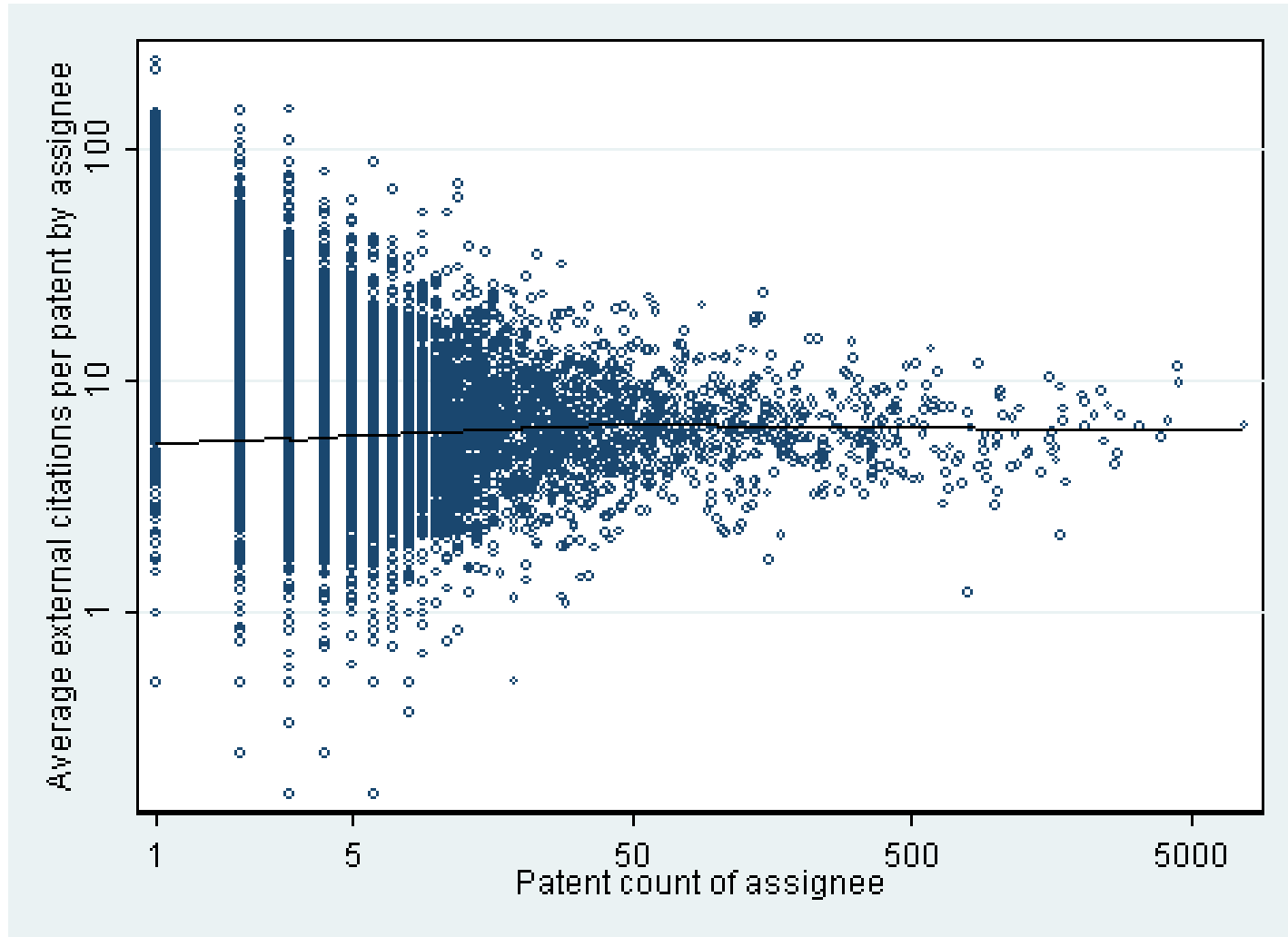
Notes: See Figure 4. Figure employs a log scale for average external citations. Figure again shows how the average number of external citations per patent does not depend strongly on assignee size conditional on being an external innovation.

Table 1: Firm size and innovation intensity among R&D producers, 1997

| | Mean R&D expenditures among R&D producers ($m) | 1997 patents among R&D producers | | R&D per employee with 1% trim ($k) | 1997 patents per employee with 1% trim | | R&D per sales with 1% trim |
		Conservative matching	Aggressive matching		Conservative matching	Aggressive matching	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Using employee counts to group firms							
1-100 employees	3.8	0.5	0.6	38.9	0.0141	0.0227	0.257
101-500 employees	4.3	1.3	1.5	22.5	0.0053	0.0063	0.128
501-1500 employees	6.1	2.4	3.6	8.0	0.0027	0.0040	0.035
1501-5000 employees	17.5	6.7	13.2	6.7	0.0026	0.0053	0.032
5001+ employees	149.0	62.8	83.2	5.8	0.0032	0.0046	0.018
B. Using establishment counts to group firms							
1 establishment	3.7	1.1	1.3	35.3	0.0114	0.0173	0.216
2 establishments	4.7	1.5	1.9	15.4	0.0054	0.0072	0.100
3-5 establishments	9.6	2.8	3.6	14.0	0.0047	0.0061	0.093
6-10 establishments	8.7	3.3	6.8	9.3	0.0022	0.0044	0.045
11-50 establishments	21.5	12.1	20.7	7.2	0.0028	0.0052	0.032
51+ establishments	144.0	54.8	69.5	5.2	0.0023	0.0030	0.018

Notes: Table reports the relationship between the firm size distribution and innovative activity for US firms undertaking R&D. Columns 1-3 report R&D expenditures and raw patent counts. Columns 4-6 report the intensity of innovation normalized by employment counts. Column 7 reports intensity of innovation normalized by sales. Larger firms undertake more R&D and obtain more patents, but conditional intensities decline with firm size. R&D expenditures are averages by firm over 1993-1997 for firms reporting R&D in 1997. Expenditures are in constant 1997 dollars. App. Tables 1a-1e report broader descriptive statistics and longitudinal changes in these distributions.

Table 2: Cross-sectional relationship between assignee size and citation behavior

	Likelihood patent has no external citation	Prevalence of patents by external citation ranks (coefficients sum to zero across columns)				Normalized likelihood of patent being among the most cited patents (likelihood normalized to have unit sd)		
		0-24%	25%-49%	50%-74%	75%-100%	Top 25%	Top 5%	Top 1%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Linear specification								
Log count of assignee patents, 1975-1984	0.003 (0.001)	0.008 (0.001)	0.001 (0.001)	-0.001 (0.001)	-0.008 (0.001)	-0.017 (0.003)	-0.007 (0.003)	-0.008 (0.004)
B. Non-parametric approach (relative to assignees with one patent)								
(0,1) Indicator variable for 2-5 patents	0.001 (0.003)	0.009 (0.004)	-0.001 (0.004)	0.002 (0.005)	-0.011 (0.005)	-0.025 (0.011)	-0.003 (0.011)	-0.006 (0.012)
(0,1) Indicator variable for 6-10 patents	0.002 (0.003)	0.015 (0.005)	0.000 (0.005)	-0.003 (0.005)	-0.012 (0.006)	-0.026 (0.014)	-0.015 (0.015)	-0.014 (0.015)
(0,1) Indicator variable for 11-20 patents	0.008 (0.004)	0.020 (0.006)	-0.008 (0.005)	-0.005 (0.006)	-0.007 (0.007)	-0.015 (0.016)	0.008 (0.016)	-0.011 (0.017)
(0,1) Indicator variable for 21-100 patents	0.010 (0.003)	0.030 (0.005)	0.001 (0.005)	-0.003 (0.005)	-0.028 (0.006)	-0.064 (0.014)	-0.030 (0.013)	-0.031 (0.013)
(0,1) Indicator variable for 101-1000 patents	0.038 (0.005)	0.058 (0.006)	0.018 (0.005)	-0.013 (0.005)	-0.063 (0.007)	-0.144 (0.017)	-0.069 (0.014)	-0.051 (0.014)
(0,1) Indicator variable for 1001+ patents	0.052 (0.011)	0.071 (0.010)	0.030 (0.009)	-0.025 (0.006)	-0.077 (0.013)	-0.175 (0.030)	-0.085 (0.020)	-0.066 (0.016)
Tech.-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	272,322	272,322	272,322	272,322	272,322	272,322	272,322	272,322

Notes: Table quantifies the relationship between patent citation behavior and assignee size. Assignee size is measured as the number of patents for the assignee. The top panel uses a linear specification, while Panel B uses a non-parametric specification. Column 1 shows that the likelihood of a patent not receiving an external citation rises with assignee size. Columns 2-5 show the distribution of external citations for the patent in quartiles by external citations received. Column 2 is the lowest quality quartile, and Column 5 is the highest quality quartile. The coefficients for a row sum to zero across these columns. Large assignees have disproportionate representation in the lowest quality quartile compared to small assignees. Columns 6-8 examine the upper quartile effect at ever increasing quality levels. The underrepresentation of large assignees weakens at increasingly high quality bars. This pattern suggests that the comparative advantage of small firms lies in making major innovations, but not necessarily breakthrough innovations. The sample includes US industrial patents applied for between 1975-1984 and their citations during 1975-1999. Estimations include technology-year fixed effects, cluster standard errors at the assignee level, and weight patents such that each assignee receives constant weight.

Table 3: Cross-sectional relationship between firm size and citation behavior

	Likelihood patent has no external citation	Prevalence of patents by external citation ranks (coefficients sum to zero across columns)				Normalized likelihood of patent being among the most cited patents (likelihood normalized to have unit sd)		
		0-24%	25%-49%	50%-74%	75%-100%	Top 25%	Top 5%	Top 1%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Linear specification								
Log firm employment, 1975-1984	0.002	0.004	-0.002	-0.001	-0.001	-0.002	0.002	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
B. Non-parametric approach (relative to firms with 100 and fewer employees)								
(0,1) Indicator variable 101-500 employees	-0.002	0.008	-0.003	0.005	-0.011	-0.025	0.005	0.030
	(0.004)	(0.006)	(0.007)	(0.007)	(0.007)	(0.016)	(0.017)	(0.020)
(0,1) Indicator variable 501-1500 employees	0.003	0.010	-0.016	0.014	-0.007	-0.017	0.012	-0.018
	(0.005)	(0.008)	(0.009)	(0.009)	(0.009)	(0.021)	(0.022)	(0.022)
(0,1) Indicator variable 1501-5000 employees	0.005	0.010	-0.009	-0.011	0.010	0.022	0.001	-0.028
	(0.006)	(0.009)	(0.010)	(0.010)	(0.010)	(0.024)	(0.024)	(0.024)
(0,1) Indicator variable 5001+ employees	0.024	0.039	-0.007	-0.020	-0.013	-0.029	0.039	0.024
	(0.006)	(0.009)	(0.009)	(0.009)	(0.010)	(0.023)	(0.028)	(0.030)
Tech.-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	257,300	257,300	257,300	257,300	257,300	257,300	257,300	257,300

Notes: See Table 2. Table considers firm size instead of assignee size. The sample of patents is reduced to those that are matched to the US Census Bureau data (conservative matching). Large firms again have disproportionate representation in the lowest quality quartile compared to small firms. The underrepresentation of large firms weakens at increasingly high quality bars. This pattern suggests that the comparative advantage of small firms lies in making major innovations, but not necessarily breakthrough innovations.

Table 4: Cross-sectional relationship of assignee size and self citation behavior

	Count of assignees by number of 1995 patents with citations for US-based, industrial patents over prior 5 years	Mean observed self citation share for the US-based, industrial patents over the prior 5 years	Comparison of observed self citation behavior against 1000 Monte Carlo simulations that replicate technologies and years of citations		
			Mean test statistic for 95% confidence level by size category	Share of firms deviating at 95% confidence level from random behavior	Mean deviations of observed citation shares (col. 2 minus col. 3)
	(1)	(2)	(3)	(4)	(5)
1 patent	8044	9%	1%	13%	8%
2-5 patents	3382	17%	3%	35%	14%
6-10 patents	595	22%	4%	64%	18%
11-20 patents	307	23%	4%	73%	19%
21-100 patents	288	27%	4%	89%	23%
101+ patents	65	31%	6%	97%	25%

Notes: Table reports the results of Monte Carlo simulations of self citation behavior by firm size. The sample is restricted to US-based, industrial patents in 1995 and their citations to other US-based, industrial patents over the prior five years. Rows group assignees by their patent counts in 1995. The second column indicates the share of observed citations that are self citations. For the Monte Carlo simulations, we draw counterfactuals that match the technologies and application years of cited patents. We include the original citation among the possible pool of patents, and we draw with replacement. We measure from the simulation a counterfactual self citation share to assignee size relationship. We repeat the simulations 1000 times to generate 95% confidence bands for the self citation ratio of each assignee. These confidence bands are specific to assignees based upon their size and underlying technologies. The third column provides the mean test statistic by firm size. This statistic rises with firm size because firms with larger patent portfolios are more likely to cite themselves even if citations are random. The fourth column indicates the share of assignees by size category that exhibit self citation behavior that exceeds a random pattern at a 95% confidence level. These deviations are strongly increasing in firm size. The last column presents the mean deviation of observed self citation behavior from the simulation baselines. These deviations are also increasing in firm size. The appendix shows that the basic self citation relationship to firm size holds within firms using panel estimation techniques.

Table 5: Panel relationship between entry and patent quality

| | Number of external citations | Number of claims on patent | Prevalence of patents by external citation ranks (coefficients sum to zero across columns) | | | |
			0-24%	25%-49%	50%-74%	75%-100%
	(1)	(2)	(3)	(4)	(5)	(6)
			A. Within assignee estimations			
First two years the assignee is observed	0.667 (0.082)	1.212 (0.140)	-0.004 (0.004)	-0.009 (0.005)	-0.005 (0.005)	0.017 (0.005)
Assignee fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Tech.-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	250,657	250,657	250,657	250,657	250,657	250,657
			B. Within firm estimations			
First two years the firm is observed	0.699 (0.107)	0.838 (0.172)	-0.012 (0.005)	-0.013 (0.006)	0.003 (0.007)	0.023 (0.006)
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Tech.-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	141,110	141,110	141,110	141,110	141,110	141,110

Notes: Table quantifies changes in average patent quality within firms over time. The top panel considers the first two years that an assignee is observed, while Panel B uses the first two years that a firm is observed. Columns 1 and 2 show that external citation rates and claims per patent are higher at firm entry. Columns 3-6 show the distribution of external citations in quartiles. Column 3 is the lowest quality quartile, and Column 6 is the highest quality quartile. The coefficients for a row sum to zero across these columns. Entrants have disproportionate representation in the highest quality quartile compared to later years for the same firm. The sample includes US industrial patents for assignees and firms first observed between 1977 and 1994. Estimations include assignee/firm fixed effects and technology-year fixed effects, cluster standard errors at the assignee/firm level, and weight patents such that each assignee/firm receives constant weight.

Table 6: Assignee size and building upon technologies

	Number of external citations on citing patent	Number of claims on citing patent	Prevalence of patents by external citation ranks among citing patents (coefficients sum to zero across columns)			
			0-24%	25%-49%	50%-74%	75%-100%
	(1)	(2)	(3)	(4)	(5)	(6)
External citation	0.849	1.236	-0.015	-0.009	-0.005	0.029
	(0.053)	(0.073)	(0.002)	(0.002)	(0.002)	(0.002)
Cited patent fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Citing tech-year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	761,940	761,940	761,940	761,940	761,940	761,940

Notes: Table characterizes differences in patent quality for internal versus external patents that cite a particular invention. Columns 1 and 2 show that external citation rates and claims are higher. Columns 3-6 show the quality distribution of the citations by quartiles. Column 3 is the lowest quality quartile, and Column 6 is the highest quality quartile. External citations are consistently of higher quality. The sample includes citations of US industrial patents from 1975-1984 applied for within ten years after the original patent. Estimations include cited patent fixed effects and technology-period fixed effects for citing patents. Estimations cluster standard errors by cited patent. The appendix extends these specifications.

Table 7: Innovation quality and employment growth of firm

	Base patent regression	Including quality distributions	Small company sample	Large company sample	Balanced panel sample	Including self citing measures
	(1)	(2)	(3)	(4)	(5)	(6)

Dependent variable is employment growth of firm

	Base patent regression	Including quality distributions	Small company sample	Large company sample	Balanced panel sample	Including self citing measures
Log employment	-0.065	-0.065	-0.115	-0.071	-0.078	-0.065
	(0.002)	(0.002)	(0.004)	(0.004)	(0.005)	(0.002)
Log patents	0.049	0.042	0.066	0.029	0.034	0.050
	(0.003)	(0.004)	(0.008)	(0.004)	(0.006)	(0.004)
Quality distribution of firm's inventions (share relative to two lowest quality quartiles):						
Third quality quartile		0.027	0.032	0.014	0.026	0.026
(50th-74th)		(0.010)	(0.012)	(0.014)	(0.023)	(0.010)
Highest quality quartile		0.049	0.057	0.032	0.081	0.048
(75th-100th)		(0.010)	(0.013)	(0.016)	(0.023)	(0.011)
Quality distribution of external inventions cited by firm (share relative to two lowest quartiles):						
Third quality quartile		0.012	0.007	0.003	0.010	0.016
(50th-74th)		(0.012)	(0.015)	(0.016)	(0.022)	(0.012)
Highest quality quartile		0.074	0.081	0.041	0.039	0.076
(75th-100th)		(0.010)	(0.012)	(0.014)	(0.019)	(0.010)
Extent to which the firm was backward self citing in patents (indicator variables relative to no self cites):						
Moderate self citation						-0.027
(>0% and <=20%)						(0.010)
High self citation						-0.074
(>20%)						(0.012)
Tech.-Period effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	29,496	29,496	18,807	10,689	7,705	29,496

Notes: Table quantifies the relationship between firm employment growth and the quality distribution of patents internal to the firm and externally cited by the firm. In particular, it shows the importance of high quality inventions (upper quartile) both within and external to the firm for subsequent employment growth. The sample includes all firms in the US Census Bureau data that are matched to patent data during the 1975-1999 period. The data center on three time periods: 1978-1982, 1983-1987, and 1988-1992. Firm growth is measured as [emp(t+1)-emp(t)]*2/[emp(t+1)+emp(t)] across these periods. The internal patent quality distribution and backwards self citation shares are determined by inventions made by the firm in the current period. The external quality distribution is determined by inventions made by other firms that are cited by the focal firm. These external inventions are limited to those made in the prior five years. Earlier and later periods than 1978-1992 are used in constructing these variables. The small company sample includes all firms that are observed to have 100 or fewer employees during the 1978-1992 period. The balanced panel sample includes all firms that patent in all three periods. Estimations include technology-period fixed effects, cluster standard errors at the firm level, and are unweighted. The appendix extends these estimations.

5 Theoretical Appendix

Proof of Proposition 1. Conjecture that $V(\mathbf{q}_f) = A\sum_{q_j \in \mathbf{q}_f} q + B\bar{q}$. Substituting this expression into the original value function

$$r^*A\sum_{q_j \in \mathbf{q}_f} q_j + r^*B\bar{q} - \bar{q}g^*B = \max_{x,[z_j]_{j \in \mathcal{J}_f}} \left\{ \begin{array}{c} \sum_{q_j \in \mathbf{q}_f} \pi^* q_j - \sum_{q_j \in \mathbf{q}_f} c_z(z_j) q_j \\ -c_x(x)\bar{q} + xA\bar{q}(1+\Gamma) \\ +\sum_{q_j \in \mathbf{q}_f} z_j A q_j \lambda - \sum_{q_j \in \mathbf{q}_f} \tau^* A q_j \end{array} \right\}.$$

This expression holds if and only if

$$r^*A = \max_z \{\pi^* - c_z(z) + zA\lambda - \tau^*A\} \tag{40}$$

$$B(r^* - g^*) = \max_x \{xA(1+\Gamma) - c_x(x)\}. \tag{41}$$

∎

Proof of Proposition 2. We will prove by induction. Define $S_n \equiv \frac{\tilde{E}^*}{E^*}\left(\frac{x^*}{\tau^*}\right)^n \frac{1}{n!}$. Recall the flow equations (26). The first two lines indicate that $S_1 = \mu_1$ and $S_2 = \mu_2$. Moreover the third equation delivers the following equality $\mu_3 = S_3$ for $n = 2$. Assume it holds up to some $n > 2$. Then the last line can be rewritten as

$$E^*\mu_{n+1}^*(n+1)\tau^* + \frac{\tilde{E}^*}{E^*}\left(\frac{x^*}{\tau^*}\right)^{n-1}\frac{E^*x^*}{(n-1)!} = E^*\frac{\tilde{E}^*}{E^*}\left(\frac{x^*}{\tau^*}\right)^n \frac{1}{n!}(x^* + n\tau^*).$$

Rearranging this equation we get

$$\mu_{n+1}^* = \frac{\tilde{E}^*}{E^*}\left(\frac{x^*}{\tau^*}\right)^{n+1}\frac{1}{(n+1)!} = S_{n+1}.$$

∎

Proof of Proposition 3. Note that $Y^* = (1-\beta)^{\frac{1-2\beta}{\beta}}\left[\frac{\varsigma}{\bar{\beta}}\right]^{\frac{1-\beta}{\beta}} L^*\bar{q}$. Therefore the growth rate of aggregate output is equivalent to the growth rate of the average quality of product lines. We can express the level of $\bar{q}(t)$ after an instant Δt as

$$\bar{q}(t + \Delta t) = \left\{ \begin{array}{c} \bar{q}(t)[\tau^*\Delta t(1+\Gamma) + z^*\Delta t(1+\lambda)] \\ +\bar{q}(t)[1 - \tau^*\Delta t - z^*\Delta t] \end{array} \right\}.$$

Now subtract $\bar{q}(t)$ from both sides and divide by Δt and take the limit as $\Delta t \to 0$

$$g = \frac{\dot{\bar{q}}(t)}{\bar{q}(t)} = \lim_{\Delta t \to 0}\frac{\bar{q}(t+\Delta t) - \bar{q}(t)}{\Delta t}\frac{1}{\bar{q}(t)} = \tau^*\Gamma + z^*\lambda.$$

∎

Proof of Proposition 4. Let Λ be a correspondence that maps $[x \quad z \quad \mathbb{S}]^T \in \Theta$ into a new vector of iterates $[x' \quad z' \quad \mathbb{S}']^T \in \Theta$ such that

$$\Lambda : [x \quad z \quad \mathbb{S}]^T \rightrightarrows [x' \quad z' \quad \mathbb{S}']^T.$$

We will define the iterations using the equations that characterize the equilibrium in Definition 3. First from (12), (13), and (29) profit can be expressed $\pi^* = (1 - \mathbb{S}^*) \beta (1 - \beta) \tilde{\beta} (1 - \beta + \beta^2)^{-1}$. Moreover from (9), (21), and (30), the equilibrium interest rate can be expressed as $r^* = \varepsilon (\mathbb{S}^* x^* \Gamma + z^* \lambda) + \rho$. As a result, from (16) the value of exploration innovation is

$$A = \frac{(1 - \mathbb{S}^*) \beta (1 - \beta) \tilde{\beta} (1 - \beta + \beta^2)^{-1} - c_z (z^*)}{\mathbb{S}^* x^* (\Gamma \varepsilon + 1) + (\varepsilon - 1) z^* \lambda + \rho}. \tag{42}$$

Using this, we can define z' as the maximizer of (16)

$$z' \left([x, z, \mathbb{S}]^T\right) = \arg \max_{\tilde{z} \in [0, \bar{z}]} \left\{ \frac{(1 - \mathbb{S}) \beta (1 - \beta) \tilde{\beta} (1 - \beta + \beta^2)^{-1} - c_z (z)}{\mathbb{S} x (\Gamma \varepsilon + 1) + (\varepsilon - 1) z \lambda + \rho} \lambda \tilde{z} - c_z (\tilde{z}) \right\}. \tag{43}$$

Similarly we use (17) to find the next iteration of x'

$$x' \left([x, z, \mathbb{S}]^T\right) = \arg \max_{\tilde{x} \in [0, \bar{x}]} \left\{ \frac{(1 - \mathbb{S}) \beta (1 - \beta) \tilde{\beta} (1 - \beta + \beta^2)^{-1} - c_z (z)}{\mathbb{S} x (\Gamma \varepsilon + 1) + (\varepsilon - 1) z \lambda + \rho} \tilde{x} (1 + \Gamma) - c_x (\tilde{x}) \right\}. \tag{44}$$

Finally we use the no-arbitrage condition (22), (17), and (42) to find the next iteration of the share of entrepreneurs,

$$\mathbb{S}' = \max \left\{ 0, \min \left\{ 1, \frac{1}{\varepsilon x \Gamma} \left[\frac{\tilde{\beta} [(\varepsilon - 1) (\mathbb{S} x \Gamma + z \lambda) + \rho]}{\frac{(1 - \mathbb{S}) \beta (1 - \beta) \tilde{\beta} (1 - \beta + \beta^2)^{-1} - c_z (z)}{\mathbb{S} x (\Gamma \varepsilon + 1) + (\varepsilon - 1) z \lambda + \rho} x (1 + \Gamma) - c_x (x)} - \rho \right] - \frac{z \lambda}{x \Gamma} \right\} \right\}. \tag{45}$$

By construction Λ maps the compact set Θ into itself. Note also that the mapping is also continuous. To finalize the proof, we need to show that $\mathbb{S} \in (0, 1)$ since we used the no-arbitrage condition with equality in (45). By contradiction assume first $\mathbb{S}^* = 0$. That implies $\tau^* = 0$, $L^* + \tilde{L}^* = 1$, $L^* = \frac{\beta (1 - \beta) \tilde{\beta}}{1 - \beta + \beta^2}$ and $r^* = z^* \lambda \varepsilon + \rho$. Using (9) and (30) to substitute for the interest rate and using the value of profit, the expected value of holding a product line (16) is $(\varepsilon z^* \lambda + \rho) A = \beta (1 - \beta) \tilde{\beta} / (1 - \beta + \beta^2) + A \lambda z^* - c_z (z^*)$. This implies

$$A = \frac{\beta (1 - \beta) \tilde{\beta} / (1 - \beta + \beta^2) - c_z (z^*)}{(\varepsilon - 1) z^* \lambda + \rho} > \frac{\beta (1 - \beta) \tilde{\beta} / (1 - \beta + \beta^2)}{(\varepsilon - 1) z^* \lambda + \rho}. \tag{46}$$

Then the expected value of entry is

$$B = \frac{\max_x \{A x (1 + \Gamma) - c_x (x)\}}{z^* \lambda (\varepsilon - 1) + \rho} > \frac{\tilde{\beta}}{z^* \lambda (\varepsilon - 1) + \rho} \underbrace{\max_x \left\{ \frac{\beta (1 - \beta) / (1 - \beta + \beta^2)}{(\varepsilon - 1) z^* \lambda + \rho} x (1 + \Gamma) - \frac{c_x (x)}{\tilde{\beta}} \right\}}_{> 1 \text{ by Assumption 1}}.$$

The equality uses the expression in (17), and the inequality follows from (46). Using Assumption 1 we get $B > \tilde{\beta} / (z^* \lambda (\varepsilon - 1) + \rho)$. Note that the value of working from (11) is $w^* / r^* = \tilde{\beta} \bar{q} / (\varepsilon z^* \lambda + \rho)$. Therefore

$$B \bar{q} > \frac{\tilde{\beta} \bar{q}}{z^* \lambda (\varepsilon - 1) + \rho} = \frac{w^*}{r^*} \frac{\varepsilon z^* \lambda + \rho}{z^* \lambda (\varepsilon - 1) + \rho} > \frac{w^*}{r^*}$$

which is a contradiction from (8).

Next assume $\mathbb{S}^* = 1$. This implies $L^* = 0$ and $\pi^* = 0$. From (43) and (44) we have $x^* = z^* = 0$. As a result $A = B = 0$. However $w^*/r^* = \tilde{\beta}\bar{q}/\rho$, which implies $w^*/r^* > B\bar{q}$. Again, this is a contradiction from (8). This completes the proof. ∎

Proof of Lemma 1. Let $j_f^n \in \mathcal{J}_f$ denote the product line in which firm f had its nth most recent exploration innovation among its current product lines. Take two firms with $n_f > n_{f'} \geq n$. Note that expected quality level of the latest nth innovation is the same across firms, $\mathbb{E}(q_{j_f^n}) = \mathbb{E}(q_{j_{f'}^n})$. To see this, consider the expected value of the latest exploration innovation of a randomly chosen firm f, $\mathbb{E}(q_{j_f^1})$. Since our analysis is on steady state, we will focus on the expected product line quality normalized by current average quality $\mathbb{E}(q_{j_f^1})/\bar{q}$ such that the expected current normalized quality is 1. Similarly, the expected initial quality of an exploration innovation that was made a periods ago is $e^{-g^* a}$. A couple of observations are in order. For an innovation to be a periods old, two events must have occurred during a periods: (i) no other firm has replaced firm f in product line j for a periods (with probability $e^{-\tau^* a}$) and (ii) from today backwards in history, the first arrival of firm f has occurred a periods ago (with probability $x^* e^{-x^* a}$). Finally, a product line's quality improves through exploitation of the same firm at the rate λz^* on average. Then the current expected quality of the latest exploration innovation of a randomly chose firm f is

$$\mathbb{E}(q_{j_f^1}) = \int_0^\infty e^{-\tau^* a} x^* e^{-x^* a} e^{-g^* a} e^{\lambda z^* a} da = x^* \int_0^\infty e^{-(\tau^* + x^*(1+\Gamma))a} da = \frac{x^*}{\tau^* + x^*(1+\Gamma)}.$$

The main conclusion of this analysis is that the expected quality is independent of the number of product lines n_f that firm f owns. Repeating the same steps for $\mathbb{E}(q_{j_f^2})$, $\mathbb{E}(q_{j_f^3})$... and so on leads to the fact that $\mathbb{E}(q_{j_f^n}) = \mathbb{E}(q_{j_{f'}^n})$ for any f, f' and n such that $n_f > n_{f'} \geq n$. Therefore

$$\mathbb{E}(Q_f \mid n_f) = \sum_{n=1}^{n_{f'}} \mathbb{E}(q_{j_f^n}) + \sum_{n=n_{f'}+1}^{n_f} \mathbb{E}(q_{j_f^n}) = \mathbb{E}(Q_{f'} \mid n_{f'}) + \sum_{n=n_{f'}+1}^{n_f} \mathbb{E}(q_{j_f^n}) > \mathbb{E}(Q_{f'} \mid n_{f'}).$$

∎

Proof of Proposition 5. Firm growth is equivalent to the growth of Q_f. After a small time interval, the quality index will be on average

$$Q_f(t+\Delta t) = \left\{ \begin{array}{c} x^* \Delta t [Q_f(t) + \bar{q}(1+\Gamma)] + \sum_{q_f} z^* \Delta t [Q_f(t) + \lambda q_f] \\ + (1 - x^* \Delta t - n_f z^* \Delta t - n_f \tau^* \Delta t) Q_f(t) \\ + \sum_{q_f} \tau^* \Delta t (Q_f - q_f) \end{array} \right\}.$$

Then after some algebra the expected growth rate of a firm is

$$\mathcal{G}(Q_f) = \lim_{\Delta t \to 0} \frac{Q_f(t+\Delta t) - Q_f(t)}{\Delta t Q_f} = \frac{x^* \bar{q}(1+\Gamma)}{Q_f} + z^* \lambda - \tau^*,$$

which is decreasing in Q_f. ∎

Proof of Proposition 6. Immediate from the text. ∎

Proof of Proposition 7. The total probability of having an innovation during Δt is $x^* \Delta t + n_f z^* \Delta t$. The probability of having a major innovation with $s_k \geq s_{\hat{k}} > \lambda$ is

$\left[1 - (1 - \theta)^{\hat{k}+1}\right] x^* \Delta t$. Then the probability of having a major innovation conditional on a successful innovation is the ratio $\left[1 - (1 - \theta)^{\hat{k}+1}\right] x^* \Delta t / \left(x^* \Delta t + n_f z^* \Delta t\right)$. ∎

Proof of Proposition 8. First we show that the value function has its form in (34). The conjecture in (34) is equivalent to (33) if and only if

$$r\tilde{A} = \max_z \left\{ \pi^* - c_z\left(z\right) + \tilde{A}\lambda z \right\}$$

and

$$rW\left(n\right) = \max_{x_n \in [0, \bar{x}]} \left\{ -nc_k\left(x_n\right) - nc_m\left(n\right) + nx_n \left[\Upsilon + W\left(n + 1\right) - W\left(n\right)\right] \right\}$$

where $\Upsilon \equiv \bar{q}\tilde{A}\left(1 + \Gamma\right)$. Recall that $c_m\left(n\right)$ stands for the convex managerial cost such that

$$c_m\left(0\right) = 0 \text{ and } c_m'\left(n\right), c_m''\left(n\right) > 0.$$

Our goal is to show that the optimal per-product-line R&D effort x_n^* is decreasing in the number of product lines n. Note that

$$c_k'\left(x_n^*\right) = W\left(n + 1\right) - W\left(n\right) + \Upsilon.$$

Therefore it is both necessary and sufficient to show that $\Delta_n \equiv W\left(n + 1\right) - W\left(n\right)$ is decreasing in n.

Before we start with the main proof, consider the following three lemmas.

Lemma 3 $W\left(n\right)$ *is bounded from above.*

Proof. Consider the per period return function $\Pi\left(x, n\right) = nx\Upsilon - nc_x\left(x\right) - nc_m\left(n\right)$. We will show first that $\Pi\left(x, n\right)$ is bounded. Define $\tilde{x} \equiv \arg\max{}_x \Pi\left(x, n\right)$. It is determined by the first order condition $\frac{\partial c_x(\tilde{x})}{\partial x} = \Upsilon$. Then $\Pi\left(x, n\right) \leq n\left[\tilde{x}\Upsilon - c_x\left(\tilde{x}\right)\right] - nc_m\left(n\right)$. Define $\tilde{n} \equiv \arg\max_{n \in \mathbb{Z}_+} \left\{n\left[\tilde{x}\Upsilon - c_x\left(\tilde{x}\right)\right] - nc_m\left(n\right)\right\}$. The existence of $\tilde{n}$ is ensured by the strict convexity of $c_m\left(\cdot\right)$. As a result we have $\Pi\left(x, n\right) \leq \Pi\left(\tilde{x}, \tilde{n}\right)$. Then we can conclude that $W\left(n\right) \leq W^{\max} \equiv \Pi\left(\tilde{x}, \tilde{n}\right)/\rho$, which is the present discounted sum of the maximum per-period returns. Note that in the last expression we used the fact that $r^* - g^* > \rho$. This completes the proof. ∎

Lemma 4 $\left\{\frac{W(n)}{n}\right\}_{n=1}^{\infty}$ *forms a decreasing sequence.*

Proof. Let us rewrite the value functions for $(n + 1)$ and n in the following way

$$\frac{r^*W\left(n + 1\right)}{n + 1} = x_{n+1}^*\Upsilon - c_k\left(x_{n+1}^*\right) - c_m\left(n + 1\right) + x_{n+1}^*\left[W\left(n + 2\right) - W\left(n + 1\right)\right]$$

$$\frac{r^*W\left(n\right)}{n} \geq x_{n+1}^*\Upsilon - c_k\left(x_{n+1}^*\right) - c_m\left(n\right) + x_{n+1}^*\left[W\left(n + 1\right) - W\left(n\right)\right].$$

Note that the second inequality follows from the fact that the argument on the right-hand side is not necessarily the optimal choice for n. Taking the difference of these two lines we get

$$r^* \left[\frac{W\left(n + 1\right)}{n + 1} - \frac{W\left(n\right)}{n}\right] \leq c_m\left(n\right) - c_m\left(n + 1\right) + x_{n+1}^*\left[\Delta_{n+1} - \Delta_n\right].$$

41

To obtain a contradiction, assume $\frac{W(N+1)}{N+1} > \frac{W(N)}{N}$ holds for some $N \in \mathbb{Z}_{++}$. Then from the last inequality and the increasing nature of $c_m(\cdot)$, we have

$$\Delta_{N+1} - \Delta_N > 0. \tag{47}$$

Note that $W(N+1) - W(N) = \frac{W(N+1)}{N+1} + N\left(\frac{W(N+1)}{N+1} - \frac{W(N)}{N}\right) > \frac{W(N+1)}{N+1}$ where the last inequality follows from the contradiction assumption. Then together with (47)

$$W(N+2) - W(N+1) > W(N+1) - W(N) > \frac{W(N+1)}{N+1}.$$

After rearranging the inequality we get

$$\frac{W(N+2)}{N+2} > \frac{W(N+1)}{N+1}.$$

Repeating the same steps we get the following generalization

$$\frac{W(n+1)}{n+1} > \frac{W(n)}{n} \text{ and } \Delta_{n+1} > \Delta_n, \text{ for } n > N.$$

However this contradicts the fact that $W(n)$ is bounded above. This completes the proof. $\blacksquare$

Lemma 5 *If* $\Delta_N > \Delta_{N-1}$ *and* $\left\{\frac{W(n)}{n}\right\}_{n=1}^{\infty}$ *is a decreasing sequence, then* $2W(N) > N(\Delta_N + \Delta_{N-1})$.

Proof. Assume $2W(N) \leq N(\Delta_N + \Delta_{N-1})$. This, together with $\Delta_N > \Delta_{N-1}$ implies $W(N) < N\Delta_N$ which in turn leads to

$$\frac{W(N)}{N} < \frac{W(N+1)}{(N+1)},$$

which is a contradiction. That completes the proof. $\blacksquare$

Proof of Proposition 8 (continued) Note that proving that x_n^* to be decreasing is equivalent to proving that $\Delta_n \equiv W(n+1) - W(n)$ is decreasing. The latter is proven next. By contradiction, assume $\exists N$ such that $\Delta_N > \Delta_{N-1}$. Next, consider the value functions

$$\frac{r^* W(N+1)}{N+1} = x_{N+1}^* \Upsilon - c_k(x_{N+1}^*) - c_m(N+1) + x_{N+1}^*[W(N+2) - W(N+1)]$$

$$\frac{r^* W(N)}{N} \geq x_{N+1}^* \Upsilon - c_k(x_{N+1}^*) - c_m(N) + x_{N+1}^*[W(N+1) - W(N)]$$

$$\frac{r^* W(N)}{N} \geq x_{N-1}^* \Upsilon - c_k(x_{N-1}^*) - c_m(N) + x_{N-1}^*[W(N+1) - W(N)]$$

$$\frac{r^* W(N-1)}{N-1} = x_{N-1}^* \Upsilon - c_k(x_{N-1}^*) - c_m(N-1) + x_{N-1}^*[W(N) - W(N-1)].$$

We subtract the second from the first and fourth from the third lines to get

$$r^*\left[\frac{W(N+1)}{N+1} - \frac{W(N)}{N}\right] \leq [c_m(N) - c_m(N+1)] + x_{N+1}^*[\Delta_{N+1} - \Delta_N]$$

$$r^*\left[\frac{W(N)}{N} - \frac{W(N-1)}{N-1}\right] \geq [c_m(N-1) - c_m(N)] + x_{N-1}^*[\Delta_N - \Delta_{N-1}].$$

We again take the difference between the two lines

$$
r^* \left(\begin{array}{c} \left[\frac{W(N+1)}{N+1} - \frac{W(N)}{N}\right] \\ -\left[\frac{W(N)}{N} - \frac{W(N-1)}{N-1}\right] \end{array} \right) \leq \underbrace{\left(\begin{array}{c} [c_m(N) - c_m(N-1)] \\ -[c_m(N+1) - c_m(N)] \end{array} \right)}_{<0} + \left(\begin{array}{c} x_{N+1}^*[\Delta_{N+1} - \Delta_N] \\ -x_{N-1}^*[\Delta_N - \Delta_{N-1}] \end{array} \right).
$$

Recall the initial assumption $\Delta_N > \Delta_{N-1}$. This implies

$$
r^* \left(\left[\frac{W(N+1)}{N+1} - \frac{W(N)}{N}\right] - \left[\frac{W(N)}{N} - \frac{W(N-1)}{N-1}\right] \right) \leq x_{N+1}^*[\Delta_{N+1} - \Delta_N].
$$

Define $\Psi \equiv \left[\frac{W(N+1)}{N+1} - \frac{W(N)}{N}\right] - \left[\frac{W(N)}{N} - \frac{W(N-1)}{N-1}\right]$. This can be also written as

$$
\Psi = \frac{N\Delta_N - W(N)}{N(N+1)} - \frac{N\Delta_{N-1} - W(N)}{N(N-1)}.
$$

After rearranging the previous expression we get

$$
\begin{aligned}
(N+1)(N-1)N\Psi &= (N-1)N\Delta_N - (N-1)W(N) - (N+1)N\Delta_{N-1} + (N+1)W(N) \\
&= N^2(\Delta_N - \Delta_{N-1}) - N(\Delta_N + \Delta_{N-1}) + 2W(N) \\
&> 0
\end{aligned}
$$

where the last inequality uses the initial contradiction assumption and Lemma 5. Therefore $\Psi > 0$ and $\Delta_{N+1} > \Delta_N$. Repeating the same steps for $n \geq N+1$ gives us the following generalization, $\Delta_{n+1} > \Delta_n$, $\forall n > N$. This result contradicts the bounded nature of $W(n)$ and completes the proof. ∎

Proof of Proposition 9. Follows from Proof of Proposition 8 and the text. ∎

Proof of Proposition 10. First we compute the number of citable patents M^*. The measure of citable patents after Δt is simply

$$
M^*(t + \Delta t) = [M^*(t) + 1](x^*\Delta t(1-\theta) + \xi z^*\Delta t) + 1 \times x^*\Delta t\theta + (1 - x^*\Delta t - \xi z^*\Delta t)M^*(t).
$$

Imposing the steady state condition $M^*(t + \Delta t) = M^*(t)$ we find $M^* = \frac{1}{\theta} + \frac{\xi z^*}{x^*\theta}$. Recall the flow equations (36) and (37). Equations (36) and (25) imply $\Upsilon_{s_k,0} = \frac{\tau^*(1-\theta)^k\theta}{M^*[\tau^*\theta + \gamma s_k(\tau^*(1-\theta) + z^*\xi)]}$. Then we can rewrite (37) in a recursive form as $\Upsilon_{s_k,n} = \Upsilon_{s_k,n-1}\frac{\gamma s_k(\tau^*(1-\theta) + z^*\xi)}{[\tau^*\theta + \gamma s_k(\tau^*(1-\theta) + z^*\xi)]}$, which implies $\Upsilon_{s_k,n} = \Upsilon_{s_k,0}\left[\frac{\gamma s_k(\tau^*(1-\theta) + z^*\xi)}{\tau^*\theta + \gamma s_k(\tau^*(1-\theta) + z^*\xi)}\right]^n$. Similar reasoning applies to $\Upsilon_{\lambda,n}$ and to the flow equations (38) and (39).

For the second part of the theorem, we just rewrite the same flow equations without the internal citations $z\xi$. Then the expressions follow. ∎

Proof of A1. Immediate from Proposition 2. ∎

Proof of A2. Let f and f' be two arbitrary firms with $n_f > n_{f'}$ product lines. The survival probability of a product line j from period t to t' is $e^{-(t'-t)\tau^*}$. Moreover, the probability that firm f will have at least one exploration innovation during the same period is $1 - e^{-(t'-t)x^*}$. Since product lines receive *iid* shocks every period, we can write the firm survival probability as

$$
\mathbf{P}_{t,t'}(f \mid n_f) = 1 - e^{-(t'-t)x^*}\prod_{j\in\mathcal{J}_f}\left(1 - e^{-(t'-t)\tau^*}\right) = 1 - e^{-(t'-t)x^*}\left(1 - e^{-(t'-t)\tau^*}\right)^{n_f}
$$

where we express the probability of a firm exit as the probability that the firm loses all of its time-t product lines $\left(1 - e^{-(t'-t)\tau^*}\right)^{n_f}$ multiplied by the probability that it does not innovate any new product during that period, $e^{-(t'-t)x^*}$. Then in all remaining events, the firm will survive. Note that $\mathbf{P}_{t,t'}\left(f \mid n_f\right) > \mathbf{P}_{t,t'}\left(f' \mid n_{f'}\right)$ iff $n_f > n_{f'}$. $\blacksquare$

Proofs of A3 and A4. Firm growth in terms of the number of product lines can be expressed as

$$g_f = \frac{n_f + m_f}{n_f} - 1 = \frac{m_f}{n_f}$$

where $m_f \in \mathbb{Z}$ is simply the net number of product innovations (new products minus the ones that are lost to other firms). Then the mean and variance of growth rate can be written as $\mathbb{E}\left(g_f\right) = \frac{1}{n_f}\mathbb{E}\left(m_f\right)$ and $\mathbb{V}\left(g_f\right) = \frac{1}{n_f^2}\mathbb{V}\left(m_f\right)$. We can express $m_f = \tilde{m}_f - \sum_{k=1}^{n_f} \iota_{f,k}$ where $\tilde{m}_f$ is the *net* number of new innovations acquired through exploration R&D and $\iota_{f,k}$ is an indicator function that becomes 1 if firm f loses its kth product line through creative destruction at the rate τ^*.

Note that the expectation and variance can be expressed as $\mathbb{E}\left(m_f\right) = \mathbb{E}\left(\tilde{m}_f\right) + \sum_{k=1}^{n_f} \mathbb{E}\left(\iota_{f,k}\right)$ and $\mathbb{V}\left(m_f\right) = \mathbb{V}\left(\tilde{m}_f\right) + \sum_{k=1}^{n_f} \mathbb{V}\left(\iota_{f,k}\right)$ since acquiring and losing product lines are independent. Since the exploration innovation process is identical for all firms $\left(x_{n_f}^* = x^*, \ \forall n_f\right)$, the expectation and variance of the net number of new innovations is the same for all firms, $\mathbb{E}\left(\tilde{m}_f\right) = \tilde{\mathbb{E}}$ and $\mathbb{V}\left(\tilde{m}_f\right) = \tilde{\mathbb{V}}, \forall f$.

Next consider the following expectations. $\mathbb{E}\left(\iota_{f,k}^2\right) = \mathbb{E}\left(\iota_{f,k}\right) = e^{-\tau^*}$, where the first equality follows from the fact that $\iota_{f,k} \in \{0,1\}$. Therefore $\mathbb{V}\left(\iota_{f,k}\right) = \mathbb{E}\left(\iota_{f,k}^2\right) - \left[\mathbb{E}\left(\iota_{f,k}\right)\right]^2 = e^{-\tau^*}\left(1 - e^{-\tau^*}\right)$. As a result, we can express the expected firm growth and variance as

$$\mathbb{E}\left(g_f\right) = \frac{\tilde{\mathbb{E}}}{n_f} - e^{-\tau^*} \text{ and } \mathbb{V}\left(g_f\right) = \frac{\tilde{\mathbb{V}}}{n_f^2} + \frac{e^{-\tau^*}\left(1 - e^{-\tau^*}\right)}{n_f}.$$

Both are decreasing in n_f. $\blacksquare$

Proof of A5. Let $a_t > a_t'$ denote the time-t ages of firms f and f' respectively. We will show that $\mathbb{E}\left(n_t - n_t' \mid a_t > a_t'\right) > 0$. Note that if the two firms had $n_\tau = n_\tau'$ for some $\tau \in [t - a_t', t]$, then the expected number of products is simply the same, $\mathbb{E}\left(n_t - n_t' \mid a_t > a_t', \tau\right) = 0$ since the numbers of products of the firms evolve independently and identically from then on. Let $\tilde{\mathbf{P}}_\tau \equiv \mathbf{P}\left(n_\tau = n_\tau' \text{ and } n_{\tau'} \neq n_{\tau'}' \ \forall \tau' \in [\tau - a_\tau', \tau]\right)$ denote the probability density that the two firms have the same number of product lines at time τ for the first time since f' entered the market in period $\tau - a_\tau'$. Right before f' entered the market, the following was true: $n_{\tau - a_\tau' - \epsilon} > n_{\tau - a_\tau' - \epsilon}'$. After the entry, if f' never catches up with the number of products of firm f with probability $\tilde{\mathbf{P}}_\tau$ in period $\tau \in [t - a_t', t]$, we have $n_t > n_t'$. If firm f catches up with firm f at some period τ, then $\mathbb{E}\left(n_t - n_t' \mid a_t > a_t', \tau\right) = 0$ as explained above. Then

$$\mathbb{E}\left(n_t - n_t' \mid a_t > a_t'\right) = \left\{ \begin{array}{l} \int_{t-a_t'}^{t} \tilde{\mathbf{P}}_\tau \underbrace{\mathbb{E}\left(n_t - n_t' \mid a_t > a_t' \text{ and } n_\tau = n_\tau'\right)}_{=0} d\tau \\[2ex] + \underbrace{\left(1 - \int_{t-a_t'}^{t} \tilde{\mathbf{P}}_\tau d\tau\right)}_{>0} \underbrace{\mathbb{E}\left(n_t - n_t' \mid a_t > a_t' \text{ and } n_\tau \neq n_\tau' \ \forall \tau \in \left[t - a_t', t\right]\right)}_{>0} \end{array} \right\}$$

$$> 0.$$

We have shown that the firm age is positively correlated with the number of products. Then the rest follows from **A2**. ∎

Proof of B1. Proven in Proposition 6. ∎

Proof of B2. For firm f, the expected number of patent counts is $x^* + n_f z^*$. Note that it is increasing in the number of product lines. Yet the intensity $x^*/n_f + z^*$ is decreasing in firm size. ∎

Proof of B3. Since firm age is shown to be positively related to n_f in the proof of **A5**, this result follows from **B1** and **B2**. ∎

Proofs of C1 and C2. They follow from Proposition 10. ∎

Proof of C3. Proposition 10 shows that $\tilde{\Upsilon}_{k,n} = \tilde{\Upsilon}_{k,0}^n \tilde{\Omega}_k$. The conditional distribution of the number of external citations (conditional on patent type) is $\tilde{\Upsilon}\left(n \mid k\right) = \left(1 - \tilde{\Omega}_k\right) \tilde{\Omega}_k^n$ for $k \in \mathbb{Z}_+$. Combining this expression with the invariant distribution in Proposition 10, the expected number of external citations to an exploration patent is simply

$$\sum_{k=0}^{\infty} \Psi_k^* \sum_{n=0}^{\infty} n \tilde{\Upsilon}\left(n \mid k\right) = \sum_{k=0}^{\infty} \theta\left(1-\theta\right)^k \frac{\tilde{\Omega}_k}{1-\tilde{\Omega}_k} = \frac{\left(1-\theta\right)\gamma\eta}{1-\alpha\left(1-\theta\right)}, \tag{48}$$

where we used the fact that $\Sigma_{n=0}^{\infty} n\Omega^n = \Omega/\left(1-\Omega\right)^2$. Similar reasoning applies to exploitation patents. The conditional citation distribution is $\tilde{\Upsilon}\left(n \mid \lambda\right) = \left(1 - \tilde{\Omega}_\lambda\right) \tilde{\Omega}_\lambda^n$. Then the expected number of external citations given to an exploitation patent is

$$\sum_{n=0}^{\infty} n \tilde{\Upsilon}\left(n \mid \lambda\right) = \sum_{n=0}^{\infty} n \left(1 - \tilde{\Omega}_\lambda\right) \tilde{\Omega}_\lambda^n = \frac{\Omega_\lambda}{1-\Omega_\lambda} = \frac{\gamma\lambda\left(1-\theta\right)}{\theta}. \tag{49}$$

Recall that $\Gamma \equiv \theta\eta/\left[1 - \alpha\left(1-\theta\right)\right]$ and $\Gamma > \lambda$. Then we get the desired result. ∎

Proof of C4. Note that the citation life of a patent ends when its technology cluster gets replaced by a new major innovation. Since it is a Poisson process, the probability of having at least one arrival of a major innovation is the remainder probability from no arrival. The probability of having no arrival of a major innovation by time t is $e^{-x\theta t}$. Therefore the probability of having at least one arrival is $\mathbf{P}\left(t\right) = 1 - e^{-x\theta t}$. ∎

Proof of D1. Probability of receiving a citation is proportional to the innovation size. Since $\Gamma > \lambda$, internal patents are less likely to be cited than external patents. Since the share of external patents decline in firm size, the result follows. ∎

Proof of D2. Self-citations come through exploitation innovations. Since the share of exploitation innovation increases in firm size, the result trivially follows. ∎

Proof of D3. This follows from the fact that average innovation size is bigger for smaller firms. Since citations (received) are proportional to the innovation size, the result follows. ∎

Proof of D4. Recall Proposition 7. The relative rate of major innovation is decreasing in firm size, $\partial \mathcal{M}\left(n\right)/\partial n = -z^* x^* \left[1 - \left(1-\theta\right)^{\hat{k}+1}\right] / \left(x^* + nz^*\right)^2$. Note that this negative relationship weakens as $\hat{k}$ gets smaller. ∎

Proofs of D5 and D6. Note that entrants start with an exploration innovation only. Then the first part of **D5** follows from **C3**. Note that we proved in **A5** that young firms are smaller in size on average. Then the second part of **D5** follows from **D3**. Since innovation at the entry is only an exploration innovation and the average citation is higher for exploration innovation, the result in **D6** follows. ∎

Proof of E1. Recall that profits, number of employees, and sales are all a linear function of the total quality Q_f. The growth rate of Q_f after an innovation of size s is simply $g_f(s) = s/Q_f$. Note that $g_f(s)$ increases in innovation size, which proves the first part. Moreover, $\partial g_f(s)/\partial Q_f = -s/Q_f^2 < 0$, which proves the second part. ∎

Proof of E2. The additional quality contributions are $\lambda \bar{q}$ and $(1+\Gamma)\bar{q}$ for an internal innovation and external innovation respectively. Since $\lambda < (1+\Gamma)$, the result follows. ∎

Proof of E3. This is true from **A1, A3, B1** and **B2.** ∎

Proof of E4. More external patents imply more number of product lines that the firm owns. Then the first part of **E4** follows from **A2.** Note that internal patents has no affect since firm survival is a function of the number of product lines. On the other hand, more external citations imply a higher number of hits received through creative destruction, which reduces the number of product lines and negatively affects firm survival. ∎

Empirical Appendix

This empirical appendix further describes our data, documents empirical extensions to our reported results, and provides additional references. The format of the appendix closely follows Section 3 for easy reference.

Data Development

Longitudinal Business Database

Each establishment in the LBD is assigned a unique, time-invariant identifier. Establishments are also linked to their parent firms in the case of multi-unit firms. For most of our analyses, we aggregate establishments up to the firm level. We typically use employment counts to measure the sizes of firms, but we find equivalent results when using the value of shipments or a count of establishments for measuring firm size. Shipments, which are not part of the LBD files, are collected from the base Economic Censuses (e.g., Census of Manufacturing, Census of Services). Jarmin and Miranda (2002) describe the construction of the LBD and provide extensive descriptive statistics. Davis et al. (1996) provide an extensive discussion of the manufacturing sector.

NSF Survey of Industrial Research and Development

The Survey of Industrial Research and Development (RAD) is the basis for publications like *Science and Engineering Indicators*, *National Patterns of R&D Resources*, and *R&D in Industry*. With appropriate clearance, researchers can access the base RAD survey responses through the Census Bureau. Surveyed firms are legally required to provide five mandatory items:

- Total costs incurred for R&D within the firm (RDTOT)

- Domestic net sales and receipts of the firm (DNS)

- Domestic net employment of the firm (DNE)

- Federally funded R&D performed within the firm (RDFED)

- State location of R&D performance (added in 2002)

In addition, a host of additional variables are collected voluntarily from large R&D performers (e.g., technology break-outs, funding by federal agency). This project focuses on the RDTOT variable, which is transformed into constant 1997 dollars and paired with employment size metrics from the LBD. Adams and Peck (1994) and Kerr and Fu (2008) provide detailed introductions to the additional, voluntary data.[27] Foster and Grim (2010) provide a comprehensive overview of the RAD and the traits of the top R&D performers in the US economy.

The most important issue for our analysis of the firm size distribution and innovation is adjustments to the RAD sampling frame. Each year, the RAD surveys with certainty the

---

[27]The voluntary data include measures of foreign R&D sourcing and domestic R&D contracting. These variables are reported for most every firm. We use these data to confirm that our firm size distribution patterns are not being driven by multinational structures, outsourcing, or similar.

identified firms that are conducting R&D within the US over a nominal expenditure bar. This expenditure hurdle began at $500,000 in the 1970s, was raised to $1 million for most of the 1980s and 1990s, was raised again to $5 million in 1996, and was most recently adjusted to $3 million in 2002. While firms undertaking less than this bar are sub-sampled, these records are more difficult to employ due to their uneven coverage and frequently imputed values. Our work concentrates on the period when the hurdle was $1 million, and we ignore values beneath this hurdle. We further comment on the expenditure hurdles when discussing the intensity measures below. The Census Bureau and NSF are constantly working to update the sample frame to include new R&D performers although short identification lags can occur.

The product versus process distinctions in Figure 1 are taken from RAD surveys between 1979 and 1989. The RAD also collected these data in 1991, but the data for this year are not reliable. The patterns shown in Figure 1 are quite robust. The increase in exploitation behavior with firm size holds with all four metrics when controlling for industry-year effects. A 10% growth in firm size is typically associated with a 0.5% growth in the likelihood of undertaking exploitation behavior and a 0.1% (non-log) share growth on the intensive margin. All of these estimates are statistically significant. No contemporaneous growth effects are evident once firm size is controlled for.[28]

Three of the four relationships unambiguously hold when further including firm fixed effects regardless of how firm size is measured. The effects on the extensive margin are comparable in size. A 10% growth in firm size is associated with a 0.3% (non-log) share growth for self citations in panel settings. The one exception is the share of R&D that is process oriented. The positive relationship of this latter metric to firm size is not evident when modelling firm size through employments in panel estimates, but it is present (0.1%) when modelling firm size through R&D expenditures.

NBER Patent Database

Our patent matching strategy builds on the work of Balasubramanian and Sivadasan (2010) and Kerr and Fu (2008). Balasubramanian and Sivadasan (2010) develop a name matching strategy that links patent assignees with firms in the LBD. This matching strategy principally pairs entities with similar names in local geographic areas—that is, it searches for firms with names similar to the patent assignee name in the city from which the patent is file. Balasubramanian and Sivadasan (2010) provide extensive descriptive statistics on the quality of their matches. This major undertaking has become the backbone for the bridge file of the NBER Patent Database to the Census Bureau data family.

We extend the Balasubramanian and Sivadasan (2010) match in two ways, although for convenience given their extensive documentation, we note that we obtain all of our results without these extensions. First, their patent file concluded with patents granted in 1999, and we extend the approach to consider patents granted up to May 2009. Our descriptive statistics run through 2000, but we use application years to group patents longitudinally. Having the additional grants that are approved during the 2000s helps ensure that we have a complete sample for the late 1990s.[29]

---

[28] The most noticeable deviation from growth invariance in Panel 1D, the 15% value for small, fast growing firms, disappears once industry-year differences are modelled.

[29] We conclude the sample in 2000 because many Census Bureau datasets (especially RAD) underwent extensive redesigns in the early 2000s and transitioned to the NAICS classification system.

Second, we incorporate a large number of additional refinements developed by Kerr and Fu (2008) and Acemoglu et al. (2010). These refinements in part come from a second, complementary name matching approach. We also work extensively to ensure that corporate subsidiaries are linked correctly to their parent firms (e.g., Johnson & Johnson files patents through many of its operating divisions). Inability to establish these linkages is the main limitation of name matching algorithms. Extensive manual efforts are undertaken to solidify these linkages for specific matching to the RAD.

Firm Size Distribution and Firm Growth Rates

The discussion in the text is complete.

Firm Size Distribution and Innovation Intensity

Our patterns in Table 1 differ somewhat from earlier work that concluded that conditional intensity was independent of size. App. Table 1a extends Table 1 to include all R&D performers during the 1975-1997 period. Column 7 of Table 1 is dropped as we do not have comparable sales data for the full period. The table reports average values with each five-year interval per firm being given equal weight. The first period of 1975-1977 contains three years. The patterns are very similar but somewhat flatter across the firm size distribution for the longer time horizon. We demonstrate below that the underlying longitudinal changes naturally link our work with earlier studies of R&D behavior and the firm size distribution.

Before examining these longitudinal shifts, we extend the firm sample from R&D performers to the universe of US firms. The broader sample is more comparable across studies. Across the complete sample, we can include the additional moments: 1) The probability of undertaking R&D or filing a patent increases with firm size and 2) The unconditional intensity of R&D expenditures or patenting by firm size increases with firm size.

The first moment is straightforward. The second moment regarding unconditional intensity is different from the conditional intensity discussed in the text. By expanding to the universe of US firms, we include the great majority of firms that do not undertake innovative activity. Unconditional intensities include zero values for all of these observations. This particularly affects the small firm sample as 98% of firms have 100 or fewer employees (compared to 67% for our innovative sample in Figure 1).

App. Table 1b documents these patterns using data from 1975 to 1997. Entries are averages over the sample period, with each five-year interval given equal weight, as discussed below. Columns 1 and 4 show that the probability that the firm conducts $1 million in R&D or is granted a patent grows with firm size. Not surprisingly, the differences across size classes are quite large. A third of all firms with more than 5000 employees are active in innovation, versus less than 0.05% in the smallest size class. The probability of innovation approaches one when looking at ever larger size categories within our 5001+ grouping.

Columns 2 and 4 show that unconditional R&D expenditures per employee and patents per employee also increase with firm size. Similar to the rising probabilities, these upward sloping intensity patterns have been found in many contexts. Finally, Columns 3 and 6 show that conditional intensities (i.e., intensity depending upon being observed in the RAD or patent data) decline in firm size. These relationships are most comparable to Table 1.

Appendix Tables 1c and 1d repeat these tabulations for five-year intervals from 1975-1997. These intervals are chosen such that the end year coincides with an Economic Census year. This segmentation aligns these descriptive statistics with the empirical estimations. The first period contains three years.

While these basic patterns are evident in each period, they are more accentuated in recent years, especially after 1987. Said differently, the higher R&D expenditures per employee conditional on performing R&D for small firms is significantly greater in 1997 than in 1977. The intensity (in constant dollars) increases seven to eight fold for the smallest firms, relative to a two-fold increase for the largest firms. Likewise, the probability of undertaking R&D or patenting declines slightly for the largest firm categories in 1997 relative to 1977; the probability doubles or triples in the smallest categories.

Unreported tabulations find that these monotonic relationships hold within each size category. The one exception to the monotonicity is that the growth in intensity among the smallest firms with ten or fewer employees (80% of all firms) is weaker than among firms with 11-100 employees. These additional breakdowns are not reported due to disclosure restrictions with the Census Bureau data. We also find similar results when segmenting firms based upon sales, although these data are only consistently available after 1992 for all sectors and thus cannot be used to discern a time trend. These relationships are evident within industries, too.

These results help reconcile recent work that finds a decreasing intensity to firm size relationship in Compustat samples (e.g., Akcigit 2010) with earlier work focused on the 1970s and 1980s that did not find a strong relationship between firm size and conditional innovation intensity (e.g., Cohen 1995, Cohen and Klepper 1992, 1996a). It is likewise interesting to note that measuring firm size through just manufacturing employment increases these elasticities to closer to one. Our work uses total firm employment.[30]

We earlier noted changes in the RAD sampling frame. We do not believe that these changes substantively influence our results. First, and most important, there is no floor for observing patent outcomes, where we find very similar results. Second, the transitions across panels in Appendix Tables 1c and 1d are smooth, which suggests that the small changes in the expenditure threshold are of limited consequence. Third, one can adjust the minimum expenditure thresholds upward for inclusion in the sample without influencing the patterns. Finally, the errors are bounded in the patent case and are very small.

App. Table 1e closes by showing that the core patenting patterns are robust across variants in metric design. The two groups of columns compare the conservative and aggressive matching strategies. The conservative matching strategy only retains the best firm match per assignee; these matches must also be above a strict quality threshold. The aggressive match employs all matches above a lower quality threshold, with multiple firm matches per assignee allowed (duplicating patents). The aggressive match results in about four times as many mappings. The central difference between the two techniques is that the aggressive strategy further accentuates relative patent contributions from small firms. This is because small firms are more likely to be eliminated when choosing the best firm match per assignee. We employ the conservative approach in our core estimations, and this approach is likely closer to a lower bound on small firm effects.

Looking down the table, Panels A and B show the sample end periods with just the end

---

[30] One potential deviation from our exploration-exploitation hypothesis is that the likelihood of conducting basic research rises with firm size.

year of patenting measured for each interval (i.e., 1997 patent for the 1993-1997 interval). Panels C and D instead use average patenting (i.e., mean patents during 1993-1997). Taking average patents across the period raises the share of firms filing patents and reduces their intensity measures. We favor the end-year measure as many firms patent infrequently. Single point measures are more representative of activity in the firm size distribution. Comparing the panels shows, however, that the central patterns are evident regardless of this design choice.

Patent Citation Behavior and Innovation Spillover Size

As simple statistics for our sample, 8.4% of all patents in this sample did not receive a citation by 1999, which is 15 or more years after the filing. If excluding self citations, the non-citation share is 10.7%. Looking at a shorter horizon of seven years after application, 22.8% of patents do not receive a citation and 27.9% do not receive an external citation. For cited patents from our 1975-1984 sample, the mean is 8.6 citations, the median is 6, the variance is 101.9, the skewness is 5.0, and the kurtosis is 59.4. The max is 280 citations versus a 99th percentile of 48 citations. Extreme values for the 1975-1984 sample are smaller than those in later periods.

The raw non-citation rate in the NBER Patent Database is 23.4%. This greater share is primarily due to limited time for patents at the end of the 1975-1999 sample period to have received citations. The raw non-citation rate through 1999 exceeds 70% for patents applied for after 1995. There is also some evidence that lower citation rates are due to a greater share of frivolous patents. Finally, non-citation rates are slightly higher for non-US patents and non-assigned patents, but these differences are small relative to the age effect.

Innovation Type and Firm Size Distribution

Exploration versus Exploitation Investments

App. Table 2 reports additional regression results for self citation behavior. These regressions employ an extended 1975-1999 sample. The self citation share is 9.2% higher for assignees with more than 1000 patents compared to assignees with 101-1000 patents (30% versus 20%). Likewise, the share is 4.5% higher for firms with more than 5000 employees compared to firms with 1500-5000 employees. These patterns hold when including assignee fixed effects and are very robust across specification variants.

Qualities of Exploration Innovations

Figure 4 uses a non-logarithmic citation scale and shows a slightly downward sloping trend line. Figure 5 uses a logarithmic citation scale instead. The relationship in Figure 5 is non-monotonic with a hump shape in the middle of the assignee size distribution. There is naturally greater variation and more extreme values among smaller assignees as the average for larger assignees is taken over more patents. But, there is also a large mass of small assignees with very low average citation rates.

As the graphs suggest, different estimation techniques yield different conclusions about the relationship between intensive patent quality and firm size. Econometric techniques stretch the citation scale in different ways and thereby emphasize different portions of the data. Across a large number of linear and non-linear specification alternatives, we find that the majority of specifications find a slight negative size relationship such that patents from larger firms receive

51

fewer external citations. But, an important number of specifications find a non-monotonic or slightly positive relationship. This sensitivity contrasts with the very robust patterns evident on the external margin. We thus conclude that external citations per patent do not depend significantly on firm size conditional on being an exploration innovation (assumption 3).

Complete Quality Distributions

App. Table 3 shows that these results hold when using claims as an alternative measure of patent value. Each patent includes a series of claims that delineate the property rights of the technology. These claims define the novel features of each invention from prior inventions and become a crucial factor in future patent infringement litigations. USPTO examiners review and modify the claims argued for by inventors in their applications, and several studies link the granted number of claims on a patent with its economic value. One again finds that the patents of large firms are disproportionately lower in quality as measured by claims. Moreover, the size differential that is pronounced at the 75th percentile is entirely eroded by the 99th percentile.[31]

Dynamic Effects

App. Table 4a separately estimates the citation specification in Table 6 by assignee size of the cited patent. The external effect is stronger in large assignees relative to smaller assignees. This is evident in both counts of raw external citations and claims and in differences across the quality distribution. App. Table 4b reports results that instead employ a breakdown by original patent quality. The external citation premium is most dramatic at the lower end of the quality distribution. These results provide suggestive evidence that internal and external innovations build differently on past work.

App. Table 5 presents the transition matrices. We examine five-year periods from 1975-1979 to 1990-1994 for the patenting by US industrial firms. Our goal is to examine and contrast the extent to which assignees move up and down the hierarchy of patent counts versus patent quality. Panel A assigns firms to quartiles based upon the average quality of their patents in each period. Panel B assigns firms to quartiles based upon their counts of patents. Panel C assigns firms to quartiles based upon the interaction of their patent count and average quality. This last measure represents a weighted impact. Rows sum to 100%.

A comparison of Panels A and B shows that firms transition across quality quartiles much faster than they do across count quartiles. Persistence in the highest quartile is only 18% for quality but 49% for patent counts. New entrants also enter higher in the quality distribution than in the count distribution. 27% of entrants enter into the highest quality quartile, compared to 16% in the count distribution. Exit is drawn evenly from the quality distribution, while greater size leads to better survival prospects in the count distribution. Panel C shows the joint consequences of these patterns in terms of weighted innovation counts. New entrants have a greater quality-weighted impact than simple patent counts would suggest. Similarly, the weighted transitions from the top quartile for incumbents are faster. These features are present in the model.

---

[31] As a methodology note, seemingly unrelated regressions deliver the same results as OLS in this context as the regressors are the same in each specification.

Innovation Type and Firm Growth Rates

App. Table 6a repeats Table 7's analysis using patent claims to measure quality distributions instead of citations. The results are qualitatively similar. For internal quality distributions, the highest quality quartile again stands out as being especially important for employment growth of firms. This is particularly true for small companies. The differences, however, are less dramatic such that moving 10% of the quality distribution from the lowest two quartiles to the highest quartile increases growth by an amount equivalent to a 6% increase in the number of patents. For external quality distributions, the primary employment growth effects come from movements out of the lowest quality quartiles. Interestingly, these effects are again concentrated in small companies.

App. Table 6b extends these employment growth results by looking at increasing thresholds of quality. We earlier demonstrated how the comparative advantage of small firms for making major innovations weakens as we examine more extreme values of the quality distribution due to the stochastic nature of innovation outcomes. The opposite prediction, however, should hold for these employment growth effects. Realized extreme values of patent quality—that is, breakthrough inventions—should prompt greater rates of growth. App. Table 6b confirms that the employment growth effects strengthen with increasing patent quality at the 95th and 99th percentiles.

Additional tests further confirm the robustness of the observed growth effects to many specification variants: dropping any sample year, focusing just on manufacturing or non-manufacturing firms, adjusting the external citation window, adjusting the aggressiveness of patent matching strategies, and similar. The results further hold individually in the great majority of sub-technology groups of the patent system. We conclude that major innovations are important for both the employment growth of the innovating firm and for the following firms that build upon the innovation.

These employment results are for continuing firms so that employment growth is consistently measured. Characterizing entry and exit is unfortunately much more difficult. For entry, we observe in the data all new firms who obtain a patent and/or hire employees. For our model, however, the relevant comparison group is those who sought to innovate and enter but were unsuccessful. This group includes some large-scale, publicized attempts (e.g., failed start-ups backed by venture capitalists) but is mostly a vast pool of unobservable efforts. The unobservable elements are especially true with respect to internal and external quality distributions. We thus need to calibrate the model to infer these properties.[32]

Exit is similarly difficult to analyze, although some traction is possible. The challenge for analyzing exit rates is that there are several processes occurring simultaneously, some within the model's scope and some outside of the model (e.g., intellectual property protections, markets for technologies). These multiple processes yield ambiguous predictions. Ceteris paribus, the strongest predictions of the model are that survival rates are increasing in firm size and in the number of contemporaneous patents obtained. This greater survival probability is due to the security of holding the leading innovation in several product lines.

App. Table 7 repeats the growth specification framework with an indicator variable for whether the firm exists in the next period or not. The results fit our basic model description

---

[32] For example, Kerr et al. (2010) describe the long tail of business plans and potential ventures seeking angel financing that would not appear in our data. These angel investors fund on the order of 5% of the proposals that they receive.

well. Survival is strongly increasing in firm size and current patent counts, which are the central predictions of the theory. There is no relationship of patent quality to survival when using claims to measure quality. On the other hand, a negative relationship to survival exists for very high quality patents when using external citations to assess quality. In both cases, building upon high quality, external innovations is associated with weak increases in survival probabilities.

Quite interestingly, these differences in the relationship of survival to how patent qualities are measured agree with the basic theory (E4). In particular, our model suggests that large numbers of external citations occur when important innovations are subsequently innovated upon by other firms. The higher share of external citations may thus be associated with cases where a major innovation preceded a lot of subsequent entry and activity (endogenously or exogenously). This interpretation is supported by the limited effect when using patent claims for measuring quality. This pattern is more in agreement with our model's basic predictions of quality invariance.

We do not want to push these results too far, especially the differences across quality measurements. As noted, there are many aspects of firm entry and exit that are outside of the scope of this framework. The central prediction that survival grows in patent counts and firm size is strongly supported, which is the most important piece for our theoretical framework. In future work, we hope to analyze the acquisition dimensions of technology markets further in the context of an endogenous growth model.

Finally, our model links major innovations to new product introductions by firms. Balasubramanian and Sivadasan (2010) document the growth in product lines among small manufacturing firms after they patent. Bernard et al. (2010) discuss the extensive product switching for US manufacturing firms following external shocks. These papers both employ the Census Bureau datasets, and we refer readers to them and to Acemoglu et al. (2010) for detailed discussions of product introductions.

App. Table 1a: Table 1 with 1975-1997 sample of R&D performers

	Mean R&D expenditures among R&D producers ($m)	Patents among R&D producers		R&D per employee with 1% trim ($k)	Patents per employee with 1% trim	
		Conservative matching	Aggressive matching		Conservative matching	Aggressive matching
	(1)	(2)	(3)	(4)	(5)	(6)
A. Using employee counts to group firms						
1-100 employees	1.9	0.3	0.4	23.2	0.0094	0.0262
101-500 employees	2.5	0.7	1.0	11.2	0.0028	0.0042
501-1500 employees	4.0	1.4	2.2	5.0	0.0015	0.0024
1501-5000 employees	10.7	3.8	6.9	4.2	0.0014	0.0026
5001+ employees	145.0	41.9	57.0	4.6	0.0018	0.0027
B. Using establishment counts to group firms						
1 establishment	2.4	0.6	0.9	19.8	0.0067	0.0168
2 establishments	2.9	0.9	1.3	8.3	0.0028	0.0061
3-5 establishments	4.8	1.4	2.0	6.6	0.0026	0.0042
6-10 establishments	5.7	2.0	3.7	5.2	0.0016	0.0028
11-50 establishments	20.6	6.0	10.1	4.8	0.0016	0.0028
51+ establishments	134.4	39.6	52.7	3.9	0.0015	0.0022

Notes: See Table 1. Table reports average values for innovative activity across full sample period of 1975-1997. Each five-year interval per firm is given equal weight. The first period of 1975-1977 contains three years. R&D expenditures are averages over time intervals in constant 1997 dollars. Patents are taken from the years closing each interval.

App. Table 1b: Firm size and innovation intensity across full sample of US firms

	(0,1) R&D>$1m in RAD survey	R&D per employee with 1% trim, all firms	R&D per employee with 1% trim, R&D>$1m	(0,1) patent	Patents per employee with 1% trim, all firms	Patents per employee with 1% trim, patents>0
	(1)	(2)	(3)	(4)	(5)	(6)
			A. Using employee counts to group firms			
1-100 employees	0.02%	0.004	17.5	0.05%	0.00009	0.1700
101-500 employees	1.4%	0.139	9.4	1.5%	0.00016	0.0104
501-1500 employees	8.1%	0.374	4.7	4.9%	0.00020	0.0040
1501-5000 employees	16.9%	0.658	3.9	11.4%	0.00028	0.0024
5001+ employees	35.9%	1.567	4.5	32.5%	0.00068	0.0021
			B. Using establishment counts to group firms			
1 establishment	0.03%	0.005	15.3	0.06%	0.00010	0.1500
2 establishments	0.3%	0.027	7.4	0.5%	0.00015	0.0308
3-5 establishments	1.1%	0.067	6.1	1.1%	0.00015	0.0138
6-10 establishments	3.0%	0.143	4.7	2.2%	0.00014	0.0062
11-50 establishments	7.9%	0.347	4.5	5.4%	0.00018	0.0034
51+ establishments	23.9%	0.905	3.8	19.8%	0.00039	0.0020

Notes: See Table 1. Table reports the relationship between the firm size distribution and innovative activity for the complete sample of US firms. Columns 1 and 4 report the extensive margin of undertaking innovative activity. Larger firms are more likely to undertake R&D or file a patent. Columns 2 and 5 report the unconditional intensity of innovation across all firms with zero values included for non-innovating firms. Unconditional intensities also rise with firm size. Finally, Columns 3 and 6 report conditional intensities among innovating firms. Conditional intensities decline with firm size. R&D expenditures are in thousands of 1997 dollars. Patent statistics use a conservative matching approach. Values are averages over the 1975-1997 period with each five-year interval receiving equal weight. Appendix Tables 1c and 1d show changes over time in these distributions.

App. Table 1c: Longitudinal employment detail for App. Table 1b

	(0,1) R&D>$1m in RAD survey	R&D per employee with 1% trim, all firms	R&D per employee with 1% trim, R&D>$1m	(0,1) patent	Patents per employee with 1% trim, all firms	Patents per employee with 1% trim, patents>0
	(1)	(2)	(3)	(4)	(5)	(6)
A. 1975-1977						
1-100 employees	0.01%	0.000	5.0	0.04%	0.00006	0.1606
101-500 employees	1.4%	0.040	2.8	1.4%	0.00012	0.0088
501-1500 employees	8.4%	0.154	1.8	5.4%	0.00022	0.0040
1501-5000 employees	18.9%	0.382	2.0	11.6%	0.00021	0.0018
5001+ employees	41.7%	1.247	3.0	37.6%	0.00065	0.0017
B. 1978-1982						
1-100 employees	0.01%	0.000	5.3	0.04%	0.00005	0.1527
101-500 employees	1.1%	0.032	2.9	1.2%	0.00010	0.0085
501-1500 employees	7.6%	0.164	2.2	4.1%	0.00014	0.0035
1501-5000 employees	15.1%	0.410	2.7	11.3%	0.00018	0.0016
5001+ employees	37.6%	1.181	3.1	33.4%	0.00048	0.0014
C. 1983-1987						
1-100 employees	0.01%	0.003	18.6	0.04%	0.00007	0.1636
101-500 employees	0.9%	0.085	9.2	1.3%	0.00012	0.0093
501-1500 employees	8.7%	0.493	5.7	4.5%	0.00014	0.0031
1501-5000 employees	16.2%	0.629	3.9	10.5%	0.00021	0.0020
5001+ employees	32.6%	1.619	5.0	30.7%	0.00048	0.0016
D. 1988-1992						
1-100 employees	0.03%	0.005	19.6	0.06%	0.00011	0.1711
101-500 employees	2.0%	0.195	9.5	1.7%	0.00017	0.0102
501-1500 employees	9.2%	0.515	5.6	4.9%	0.00019	0.0038
1501-5000 employees	18.2%	0.789	4.3	11.2%	0.00029	0.0026
5001+ employees	34.5%	1.855	5.4	30.0%	0.00061	0.0021
E. 1993-1997						
1-100 employees	0.03%	0.010	38.9	0.09%	0.00018	0.2021
101-500 employees	1.5%	0.343	22.5	1.9%	0.00028	0.0150
501-1500 employees	6.8%	0.545	8.0	5.4%	0.00029	0.0055
1501-5000 employees	16.1%	1.078	6.7	12.6%	0.00051	0.0040
5001+ employees	33.2%	1.932	5.8	31.0%	0.00116	0.0037

Notes: See App. Table 1b.

App. Table 1d: Longitudinal establishment detail for App. Table 1b

	(0,1) R&D>$1m in RAD survey	R&D per employee with 1% trim, all firms	R&D per employee with 1% trim, R&D>$1m	(0,1) patent	Patents per employee with 1% trim, all firms	Patents per employee with 1% trim, patents>0
	(1)	(2)	(3)	(4)	(5)	(6)

			A. 1975-1977			
1 establishment	0.02%	0.001	3.6	0.04%	0.00007	0.1472
2 establishments	0.3%	0.010	3.2	0.5%	0.00017	0.0339
3-5 establishments	1.1%	0.027	2.5	1.0%	0.00015	0.0141
6-10 establishments	3.1%	0.068	2.2	2.4%	0.00013	0.0054
11-50 establishments	8.7%	0.186	2.1	6.0%	0.00018	0.0031
51+ establishments	24.9%	0.674	2.7	22.4%	0.00039	0.0017
			B. 1978-1982			
1 establishment	0.01%	0.001	5.0	0.04%	0.00005	0.1324
2 establishments	0.3%	0.007	2.3	0.4%	0.00015	0.0348
3-5 establishments	1.0%	0.024	2.4	0.9%	0.00010	0.0105
6-10 establishments	2.7%	0.065	2.4	1.7%	0.00010	0.0058
11-50 establishments	7.7%	0.184	2.4	5.3%	0.00015	0.0027
51+ establishments	25.1%	0.735	2.9	20.6%	0.00032	0.0016
			C. 1983-1987			
1 establishment	0.03%	0.004	14.8	0.05%	0.00007	0.1446
2 establishments	0.2%	0.017	8.7	0.4%	0.00013	0.0300
3-5 establishments	0.9%	0.053	5.7	1.1%	0.00014	0.0128
6-10 establishments	2.8%	0.108	3.9	2.1%	0.00011	0.0054
11-50 establishments	8.1%	0.452	5.6	5.6%	0.00015	0.0026
51+ establishments	23.6%	0.922	3.9	19.5%	0.00031	0.0016
			D. 1988-1992			
1 establishment	0.04%	0.007	17.8	0.07%	0.00011	0.1530
2 establishments	0.5%	0.040	7.5	0.5%	0.00014	0.0259
3-5 establishments	1.4%	0.077	5.7	1.1%	0.00015	0.0138
6-10 establishments	3.3%	0.178	5.5	2.2%	0.00012	0.0053
11-50 establishments	8.1%	0.411	5.1	5.0%	0.00017	0.0035
51+ establishments	24.5%	1.074	4.4	18.5%	0.00037	0.0020
			E. 1993-1997			
1 establishment	0.04%	0.014	35.3	0.1%	0.00018	0.1727
2 establishments	0.4%	0.062	15.4	0.6%	0.00018	0.0295
3-5 establishments	1.1%	0.152	14.0	1.2%	0.00022	0.0178
6-10 establishments	3.2%	0.295	9.3	2.6%	0.00023	0.0089
11-50 establishments	7.0%	0.503	7.2	5.3%	0.00026	0.0049
51+ establishments	21.4%	1.118	5.2	18.1%	0.00055	0.0030

Notes: See App. Table 1b.

App. Table 1e: Comparison of matching strategies

	Conservative matching approach			Aggressive matching approach		
	(0,1) patent	Patents per employee with 1% trim, all firms	Patents per employee with 1% trim, patents>0	(0,1) patent	Patents per employee with 1% trim, all firms	Patents per employee with 1% trim, patents>0
	(1)	(2)	(3)	(4)	(5)	(6)
A. 1975-1977 with 1977 patents						
1-100 employees	0.04%	0.00006	0.1606	0.08%	0.00077	0.9290
101-500 employees	1.4%	0.00012	0.0088	2.5%	0.00067	0.0266
501-1500 employees	5.4%	0.00022	0.0040	8.8%	0.00065	0.0074
1501-5000 employees	11.6%	0.00021	0.0018	18.7%	0.00049	0.0026
5001+ employees	37.6%	0.00065	0.0017	48.5%	0.00129	0.0027
B. 1993-1997 with 1997 patents						
1-100 employees	0.09%	0.00018	0.2021	0.13%	0.00109	0.8377
101-500 employees	1.9%	0.00028	0.0150	2.8%	0.00074	0.0267
501-1500 employees	5.4%	0.00029	0.0055	7.5%	0.00061	0.0081
1501-5000 employees	12.6%	0.00051	0.0040	16.9%	0.00114	0.0067
5001+ employees	31.0%	0.00116	0.0037	38.2%	0.00183	0.0048
C. 1975-1977 with average patents per interval						
1-100 employees	0.10%	0.00007	0.0685	0.17%	0.00061	0.3559
101-500 employees	2.8%	0.00012	0.0045	4.9%	0.00070	0.0144
501-1500 employees	8.8%	0.00022	0.0025	13.4%	0.00064	0.0048
1501-5000 employees	18.0%	0.00022	0.0012	26.9%	0.00052	0.0019
5001+ employees	44.3%	0.00065	0.0015	55.6%	0.00131	0.0023
D. 1993-1997 with average patents per interval						
1-100 employees	0.20%	0.00013	0.0637	0.29%	0.00054	0.1799
101-500 employees	3.8%	0.00023	0.0061	5.5%	0.00057	0.0105
501-1500 employees	9.3%	0.00023	0.0025	12.5%	0.00046	0.0037
1501-5000 employees	18.6%	0.00040	0.0021	23.6%	0.00084	0.0036
5001+ employees	39.6%	0.00087	0.0022	47.3%	0.00138	0.0029

Notes: See App. Table 1b.

App. Table 2: Firm size and self citation behavior

	Fraction of current citations that are self citations				Fraction of current citations that are self citations		
	(1)	(2)	(3)		(4)	(5)	(6)
Log count of assignee patents during period	0.057 (0.001)		0.017 (0.002)	Log firm employment during period	0.009 (0.001)		0.016 (0.002)
(0,1) Indicator variable for 2-5 patents		0.065 (0.001)		(0,1) Indicator variable 101-500 employees		0.027 (0.003)	
(0,1) Indicator variable for 6-10 patents		0.114 (0.004)		(0,1) Indicator variable 501-1500 employees		0.022 (0.004)	
(0,1) Indicator variable for 11-20 patents		0.126 (0.005)		(0,1) Indicator variable 1501-5000 employees		0.031 (0.004)	
(0,1) Indicator variable for 21-100 patents		0.204 (0.007)		(0,1) Indicator variable 5001+ employees		0.076 (0.005)	
(0,1) Indicator variable for 101+ patents		0.296 (0.017)					
Year fixed effects	Yes	Yes	Yes	Year fixed effects	Yes	Yes	Yes
Assignee fixed effects			Yes	Firm fixed effects			Yes
Observations	192,607	192,607	192,607	Observations	92,902	92,902	92,902

Notes: See Table 4. Table quantifies the relationship between self citation behavior and firm size. Columns 1-3 consider assignee sizes, while Columns 4-6 consider firm sizes. The sample includes US industrial patents during the 1975-1999 period. Patents and self citation behavior are aggregated to the assignee and firm levels by year for estimation. Estimations cluster standard errors at the assignee/firm level and are unweighted.

App. Table 3: Table 2 using claims to measure patent quality

	Number of claims on patent	Prevalence of patents by claims ranks (coefficients sum to zero across columns)				Normalized likelihood of patent being among highest claim patents (likelihood normalized to have unit sd)		
		0-24%	25%-49%	50%-74%	75%-100%	Top 25%	Top 5%	Top 1%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(0,1) Indicator variable for 2-5 patents	0.233 (0.109)	-0.005 (0.004)	0.001 (0.004)	-0.007 (0.005)	0.011 (0.005)	0.024 (0.011)	0.016 (0.011)	0.016 (0.011)
(0,1) Indicator variable for 6-10 patents	0.689 (0.157)	-0.015 (0.005)	-0.005 (0.005)	-0.006 (0.005)	0.025 (0.007)	0.058 (0.015)	0.057 (0.015)	0.049 (0.015)
(0,1) Indicator variable for 11-20 patents	0.734 (0.186)	-0.021 (0.006)	0.003 (0.005)	-0.012 (0.006)	0.030 (0.007)	0.067 (0.017)	0.039 (0.017)	0.048 (0.018)
(0,1) Indicator variable for 21-100 patents	0.323 (0.173)	-0.014 (0.006)	0.001 (0.005)	-0.002 (0.005)	0.015 (0.007)	0.034 (0.017)	0.025 (0.015)	0.026 (0.013)
(0,1) Indicator variable for 101-1000 patents	-0.347 (0.228)	0.011 (0.008)	0.011 (0.005)	-0.006 (0.006)	-0.015 (0.009)	-0.034 (0.020)	-0.013 (0.017)	0.002 (0.013)
(0,1) Indicator variable for 1001+ patents	-1.170 (0.406)	0.032 (0.016)	0.034 (0.007)	-0.007 (0.009)	-0.059 (0.014)	-0.133 (0.031)	-0.069 (0.027)	-0.020 (0.020)
Tech.-Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	272,322	272,322	272,322	272,322	272,322	272,322	272,322	272,322

Notes: See Table 2.

App. Table 4a: Table 6 with firm size breakdowns

	Number of external citations on citing patent	Number of claims on citing patent	Prevalence of patents by external citation ranks among citing patents (coefficients sum to zero across columns)			
			0-24%	25%-49%	50%-74%	75%-100%
	(1)	(2)	(3)	(4)	(5)	(6)
A. Full sample of citations						
External citation	0.849	1.236	-0.015	-0.009	-0.005	0.029
	(0.053)	(0.073)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	761,940	761,940	761,940	761,940	761,940	761,940
B. Restricted to cited assignees with 1-100 patents						
External citation	0.217	0.170	0.004	-0.002	-0.007	0.005
	(0.107)	(0.146)	(0.003)	(0.004)	(0.004)	(0.005)
Observations	277,708	277,708	277,708	277,708	277,708	277,708
C. Restricted to cited assignees with 101-1000 patents						
External citation	0.829	0.939	-0.014	-0.003	-0.004	0.020
	(0.080)	(0.122)	(0.003)	(0.003)	(0.004)	(0.004)
Observations	248,264	248,264	248,264	248,264	248,264	248,264
D. Restricted to cited assignees with more than 1000 patents						
External citation	1.202	2.083	-0.026	-0.019	-0.004	0.049
	(0.087)	(0.115)	(0.003)	(0.003)	(0.003)	(0.004)
Observations	235,968	235,968	235,968	235,968	235,968	235,968
Cited patent fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Citing tech-year effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: See Table 6.

App. Table 4b: Table 6 with cited patent quality breakdowns

	Number of external citations on citing patent	Number of claims on citing patent	Prevalence of patents by external citation ranks among citing patents (coefficients sum to zero across columns)			
			0-24%	25%-49%	50%-74%	75%-100%
	(1)	(2)	(3)	(4)	(5)	(6)

A. Full sample of citations						
External citation	0.849	1.236	-0.015	-0.009	-0.005	0.029
	(0.053)	(0.073)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	761,940	761,940	761,940	761,940	761,940	761,940
B. Lowest quartile of cited patent quality						
External citation	1.786	1.449	-0.106	-0.034	0.061	0.079
	(0.202)	(0.381)	(0.014)	(0.015)	(0.013)	(0.012)
Observations	40,026	40,026	40,026	40,026	40,026	40,026
C. Second lowest quartile of cited patent quality						
External citation	1.321	1.054	-0.050	-0.030	0.016	0.064
	(0.109)	(0.207)	(0.006)	(0.007)	(0.007)	(0.006)
Observations	96,042	96,042	96,042	96,042	96,042	96,042
D. Second highest quartile of cited patent quality						
External citation	1.282	1.310	-0.026	-0.026	-0.009	0.060
	(0.081)	(0.134)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	189,139	189,139	189,139	189,139	189,139	189,139
E. Highest quartile of cited patent quality						
External citation	0.461	1.230	0.003	0.004	-0.010	0.004
	(0.075)	(0.097)	(0.002)	(0.002)	(0.003)	(0.003)
Observations	436,733	436,733	436,733	436,733	436,733	436,733
Cited patent fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Citing tech-period effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: See Table 6.

App. Table 5: Transition matrices for patent quality and quantity

| | Assignee position in the following period: | | | |
	Not present	Bottom two quartiles	Second highest quartile	Highest quartile
	(1)	(2)	(3)	(4)

A. Quality distribution of patents

Assignee position in current period:				
Not present		51%	23%	27%
Bottom two quartiles	64%	23%	9%	4%
Second highest quartile	56%	18%	17%	9%
Highest quartile	61%	9%	11%	18%

B. Count distribution of patents

Assignee position in current period:				
Not present		68%	17%	16%
Bottom two quartiles	76%	12%	5%	7%
Second highest quartile	59%	16%	9%	17%
Highest quartile	31%	12%	9%	49%

C. Weighted distribution of patents

Assignee position in current period:				
Not present		55%	28%	18%
Bottom two quartiles	73%	14%	8%	5%
Second highest quartile	63%	14%	12%	12%
Highest quartile	36%	9%	12%	43%

Notes: Table documents transition matrices across five-year periods from 1975-1979 to 1990-1994 for the patenting by US industrial firms. Rows sum to 100%. Panel A groups assignees to quartiles based upon the average quality of their patents in each period. Panel B groups assignees to quartiles based upon their counts of patents. Panel C groups assignees to quartiles based upon the interaction of their patent count and average quality. Assignees transition across quality quartiles much faster than they do across count quartiles. New entrants also enter higher in the quality distribution than in the count distribution. New entrants thus have a greater quality-weighted impact, and weighted transitions from the top quartile for incumbents are faster than simple patent counts suggest.

App. Table 6a: Table 7 using claims to measure patent quality

	Base patent regression	Including quality distributions	Small company sample	Large company sample	Balanced panel sample	Including self citing measures
	(1)	(2)	(3)	(4)	(5)	(6)
			Dependent variable is employment growth of firm			
Log employment	-0.065	-0.065	-0.115	-0.071	-0.078	-0.066
	(0.002)	(0.002)	(0.004)	(0.004)	(0.005)	(0.002)
Log patents	0.049	0.047	0.073	0.031	0.036	0.055
	(0.003)	(0.004)	(0.008)	(0.004)	(0.006)	(0.004)
Quality distribution of firm's inventions (share relative to two lowest quality quartiles):						
Third quality quartile		0.008	0.007	0.003	0.028	0.009
(50th-74th)		(0.010)	(0.012)	(0.015)	(0.023)	(0.010)
Highest quality quartile		0.030	0.035	0.002	0.032	0.029
(75th-100th)		(0.010)	(0.012)	(0.014)	(0.021)	(0.010)
Quality distribution of external inventions cited by firm (share relative to two lowest quartiles):						
Third quality quartile		0.023	0.023	0.017	0.015	0.029
(50th-74th)		(0.012)	(0.015)	(0.017)	(0.022)	(0.010)
Highest quality quartile		0.030	0.032	0.016	-0.004	0.038
(75th-100th)		(0.011)	(0.015)	(0.016)	(0.022)	(0.011)
Extent to which the firm was backward self citing in patents (indicator variables relative to no self cites):						
Moderate self citation						-0.026
(>0% and <=20%)						(0.010)
High self citation						-0.083
(>20%)						(0.012)
Tech.-Period effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	29,496	29,496	18,807	10,689	7,705	29,496

Notes: See Table 7. Table substitutes quality distributions measured through claims on patents.

App. Table 6b: Table 7 with higher thresholds

	Using external citations to measure quality			Using claims to measure quality		
	75th percentile and higher	95th percentile and higher	99th percentile and higher	75th percentile and higher	95th percentile and higher	99th percentile and higher
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable is employment growth of firm					
Log employment	-0.065	-0.065	-0.065	-0.065	-0.065	-0.065
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Log patents	0.043	0.046	0.048	0.048	0.049	0.049
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Share of own inventions in indicated quality range	0.039	0.084	0.156	0.028	0.046	0.053
	(0.010)	(0.022)	(0.054)	(0.009)	(0.019)	(0.048)
Share of cited inventions in indicated quality range	0.075	0.114	0.154	0.028	0.026	0.021
	(0.010)	(0.017)	(0.034)	(0.011)	(0.024)	(0.055)
Tech.-Period fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	29,496	29,496	29,496	29,496	29,496	29,496

Notes: See Table 7 and App. Table 6a. Specifications examine the upper quartile effect of patent quality on firm employment growth at ever increasing quality levels. The effects increase as patent quality thresholds are raised.

App. Table 7: Survival analysis

	Base survival regression	Firm survival analysis using patent claims to measure quality	Firm survival analysis using external citations to measure quality
	(1)	(2)	(3)
Log employment	0.029	0.029	0.028
	(0.001)	(0.001)	(0.001)
Log patents	0.006	0.006	0.004
	(0.001)	(0.001)	(0.001)
Quality distribution of firm's inventions (share relative to two lowest quality quartiles):			
Third quality quartile		0.000	-0.011
(50th-74th)		(0.005)	(0.005)
Highest quality quartile		0.001	-0.039
(75th-100th)		(0.005)	(0.005)
Quality distribution of external inventions cited by firm (share relative to two lowest quartiles):			
Third quality quartile		0.007	0.016
(50th-74th)		(0.006)	(0.006)
Highest quality quartile		0.008	0.020
(75th-100th)		(0.005)	(0.005)
Tech.-Period fixed effects	Yes	Yes	Yes
Observations	34,157	34,157	34,157

Notes: See Table 7. Estimations consider whether firms survive to the following period or not. The central prediction of the model is that larger firms and firms with greater numbers of contemporaneous patents survive longer. These effects are evident in the first two rows. The model delivers ambiguous predictions regarding survival and patent quality. The model predicts that survival is invariant to general patent quality. This is most evident in the second column using patent claims to measure quality. The model also predicts that high external citations are associated with greater replacement even though patent quality is high. This is evident in the third column. The appendix text further discusses this analysis.