NBER WORKING PAPER SERIES

YOUNG, RESTLESS AND CREATIVE: OPENNESS TO DISRUPTION AND CREATIVE INNOVATIONS

Daron Acemoglu
Ufuk Akcigit
Murat Alp Celik

Working Paper 19894
http://www.nber.org/papers/w19894

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2014

We thank Olga Denislamova, Hyunjin Kim, and Gokhan Oz for excellent research assistance, and conference and seminar participants at the NBER Productivity group, CIFAR, Brown University, Harvard, MIT, University of Pennsylvania, 2014 AEA Meetings and the Toulouse Network on Information Technology and Joshua Gans, Samuel Kortum, Ben Pugsley, and Pascual Restrepo for helpful suggestions and comments. We also thank the Bilkent University Economics Department for the great hospitality during this project. Financial support from the National Science Foundation, the Alfred P. Sloan Foundation, the Ewing Marion Kauffman Foundation, ARO MURI W911NF-12-1-0509 and from the Toulouse Network on Information Technology is gratefully acknowledged. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Young, Restless and Creative: Openness to Disruption and Creative Innovations
Daron Acemoglu, Ufuk Akcigit, and Murat Alp Celik
NBER Working Paper No. 19894
February 2014, Revised August 2015
JEL No. O33,O40,O43,P10,P16,Z1

ABSTRACT

This paper argues that openness to new, unconventional and disruptive ideas has a first-order impact on creative innovations—innovations that break new ground in terms of knowledge creation. After presenting a motivating model focusing on the choice between incremental and radical innovation, and on how managers of different ages and human capital are sorted across different firms with different degrees of openness to disruption, we provide firm-level, patent-level and cross-country evidence consistent with this pattern. Our measures of creative innovations proxy for innovation quality (average number of citations per patent) and creativity (fraction of superstar innovators, the likelihood of a very high number of citations, and generality of patents). Our main proxy for openness to disruption is the age of the manager—based on the idea that only companies or societies open to such disruption will allow the young to rise up within the hierarchy. Using this proxy at the firm, patent and country level, we present robust evidence that openness to disruption is associated with more creative innovations, but we also show that once the effect of the sorting of young managers to firms that are more open to disruption is factored in, the (causal) impact of manager age on creative innovations is small.

Daron Acemoglu
Department of Economics, E18-269D
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and CIFAR
and also NBER
daron@mit.edu

Ufuk Akcigit
Department of Economics
University of Chicago
1126 East 59th Street
Saieh Hall, Office 403
Chicago, IL 60637
and NBER
uakcigit@uchicago.edu

Murat Alp Celik
3718 Locust Walk
McNeil 160
Philadelphia, PA, 19104
celik@sas.upenn.edu

1 Introduction

This paper investigates the impact of economic and social incentives on "creative innovations," which we identify with the most influential, innovative and original patents. Though there are currently more than half a million patents granted by the US Patent and Trademark Office (USPTO) per year, only a handful are truly transformative in terms of their contribution to society's knowledge and their impact on the organization of production, and probably only a small fraction account for the bulk of the value created.[1] For example, within the field of drugs and medical inventions, there were 223,452 patents between 1975 and 2001, but the median number of citations of these patents within the next five years was four. A few patents receive many more citations, however. One was the patent for "systems and methods for selective electrosurgical treatment of body structures" by the ArthroCare Corporation (with 50 citations), which has also had a major impact on the field by improving many existing surgical procedures and devices used, inter alia, in arthroscopy, neurology, cosmetics, urology, gynecology and laparoscopy/general surgery. Another example is Amazon's patent for "method and system for placing a purchase order via a communications network," which received 263 citations within five years (while the median number of citations within this class is five) and has fundamentally altered online businesses.

An idea dating back to Joseph Schumpeter (1934) associates creative innovations and entrepreneurship not only with economic rewards to this type of transformative idea, but also with the ability and desire of potential innovators and entrepreneurs to significantly deviate from existing technologies, practices and rules of organization and society and to engage in "disruptive innovations." This is natural; as Schumpeter emphasizes, innovation is a deviation from existing, inertial ways of doing things, and thus relies on "mental freedom" from, or even "rebellion" against, the status quo (pp. 86-94). Similarly, technologies that will cause the most fundamental "creative destruction" naturally correspond to, and perhaps are driven by, "deviant" and disruptive behavior. This notion is pithily captured by an inscription prominently displayed on the walls of Facebook's headquarters in Silicon Valley:

> "Move fast and break things."

This perspective suggests that societies and organizations that impose a set of rigidly specified rules, discourage initiative and deviations from established norms, shun or even ostracize rebellious behavior, and do not tolerate those that "move fast and break things" will significantly lag behind their more open, "individualistic" or "risk-taking" counterparts in creative innovations—even though they might still be able to function successfully with existing technologies. In the rest of the

[1] See, among others, Trajtenberg (1990), Harhoff et al. (1999) and Sampat and Ziedonis (2004) on the relationship between citations and patent quality.

paper, we thus refer to this constellation of social and economic incentives as *openness to disruption* (short for openness to disruptive innovations, ideas and practices).

We first provide a simple model of the interplay between "*corporate culture*" (firm type capturing how open the firm is to disruption) and innovation strategies. Firms can engage in an *incremental innovation* by building on their existing leading-edge products. In addition, high-type firms (those that have a corporate culture open to disruption) can attempt a *radical innovation*, which involves combining diverse ideas to generate a technological improvement in a new area. We also assume that the skills of young managers who have more recently acquired general skills (or are less beholden to a particular type of product or technology) can be fruitfully utilized in the process of radical innovation. In the model, though incremental innovations also increase productivity, it is the radical innovations that are the engine of growth. This is because incremental innovations in a particular "technology cluster" run into diminishing returns (as in Akcigit and Kerr, 2010, or Abrams et al., 2013), while radical innovations create new technology clusters, which increase productivity directly, and also indirectly, by making another series of incremental innovations possible.

Our model predicts a reduced-form cross-sectional relationship between manager age and radical innovation. But this relationship does not correspond to the causal effect of manager age on creative innovations. Rather, manager age is both an economically relevant variable and more generally a proxy for openness to disruption, as highlighted by our model where young managers tend to work in firms that are open to radical innovation, but also contribute to the likelihood of radical innovations in such firms. These forces can also be seen from the longitudinal predictions of the model: firms that hire younger managers should subsequently have more creative innovations (because hiring a young manager is associated either with a change in a firm's type or a change in the firm's innovation strategy as it runs out of productive incremental innovation opportunities). But because firms that are more open to disruption need not immediately hire a young manager, the increase in creative innovations can precede the hiring of a younger manager.

The model further clarifies that radical innovations will generate higher quality patents that are more likely to receive a high number of citations and tend to be more general in terms of the range of citations they receive (because they are expanding into new areas), and this provides us with an empirical strategy to measure the creativity of innovations (and present evidence about several aspects of the model's implications).

Our theoretical framework also predicts another relationship we investigate empirically: products with higher sales will encourage even high-type firms and young managers to pursue incremental innovations (because of Arrow's (1962) *replacement effect*), and those with many patents will tilt things in favor of radical innovations (because of diminishing returns and more generally because there is a substantial knowledge base to build upon for such an expansion).

Finally, our model further suggests that institutions or attitudes that ban or discourage expansion into new areas or combinations that have not been previously experimented with can be highly detrimental to radical innovations. Equally, those that prevent young managers from leading companies could slow down creative innovations by failing to use their more recent vintage skills in radical innovations. Such institutions and attitudes typically vary across countries, and this reasoning suggests that similar relationships might be found in the cross-country data.

The bulk of our paper comprises an empirical study of the ideas illustrated by our theoretical model. We investigate whether companies with younger managers engage in more radical and creative innovations. As already noted, manager age is a natural proxy for openness to disruption, since companies with a corporate culture open to disruption are more likely to allow young managers to rise up to the top of the corporate hierarchy.[2]

Our empirical work uses several different measures of creative innovations, all computed from the USPTO data. These are the *average number of citations per patent*; *the fraction of superstar innovators*, which corresponds to the fraction of patents accruing to an innovator classified as a "superstar" on the basis of the number of citations; *tail innovations*, which we measure as the fraction of patents (of a country or company) that are at the pth percentile of the overall citations distribution (such as the 99th percentile) relative to those that are at the median, thus capturing the likelihood of receiving a very high number of citations normalized by the "median" number of citations; and *generality index*, constructed by Hall, Jaffe and Trajtenberg (2001), which measures the dispersion of the citations that a patent receives from different technology classes. We report several salient and robust pattern using these data.

First, we establish a very robust cross-sectional correlation between CEO (or top management) age and all of our measures of firm-level creative innovation (with or without a variety of firm-level controls). In summary, firms that tend to employ younger CEOs receive a greater number of citations per patent, have a greater fraction of their patents generated by superstar innovators, have more tail innovations, which are at the very high percentiles of the citations distribution, and have more general patents.

Second, we find similar (but somewhat smaller) results when we focus on "within-firm" variation generated by CEO changes: when a younger CEO takes charge, innovations (new patent applications) become more creative. Recall, however, that, as our theoretical analysis highlights, these within-firm results are still a mixture of the sorting effects and the causal effect of manager age on creative innovations.

Third, related to this last point and again consistent with our theoretical model, we show that

[2] Interestingly, in the examples of major innovations mentioned above, these were produced by companies with unusually young leadership. The average age of top managers at ArthroCare Corporation was 41 at the time, and only 33 at Amazon (compared to an average age of 54.84 among Compustat companies).

there is a significant increase in creative innovations *before* a firm switches to a younger CEO, but once it does make the switch, there is a further increase in the creativity of their innovations.

Fourth, we also use the structure of our model, in conjunction with the reduced-form patterns in the data, to shed further light on the relative roles of sorting and the causal effect of manager age on innovation. Namely, we utilize a simple indirect inference procedure to estimate from the reduced-form regression coefficients some of the key parameters of our model, including those governing the causal effect of manager age on creative innovations. This exercise implies that the causal effect of manager age on creative innovations is small and is dwarfed by the sorting effects resulting from the fact that firms that are more open to disruption, and thus more creative, tend to hire younger managers.

Fifth, we exploit the patent-level variation to estimate the separate impacts of CEO and inventor age on the creativity of innovations. Our results indicate that both matter, with roughly similar magnitudes. But we also find that younger CEOs tend to work with younger inventors (though CEO age has a fairly precisely estimated impact even after controlling for inventor age). These two findings, which suggest that firms typically undertake many associated changes while they are switching towards generating more creative innovations, further corroborate our interpretation that much of the cross-sectional (and within-firm) evidence reflects sorting of younger managers to firms with corporate cultures that are more open to disruption.

Finally, we also use the firm-level data to shed light on our model's prediction that firms with greater sales should be less willing to encourage new, potentially disruptive ideas, practices and innovations, while firms that are technologically more advanced, and thus not able to profitably function without engaging in major innovations, should be more likely to encourage this type of disruptive innovation. Our firm-level data enable us to investigate this idea by simultaneously including interactions of CEO age with (log) sales and (log) number of patents of the firm. Though the results here are somewhat less strong than our main findings, they are broadly consistent with the notion that CEO age interacts negatively with sales and positively with the number of patents.

We conclude the paper by showing in Section 5 that the firm-level results aggregate up to the country level, so that countries that employ younger managers appear to have more creative innovations controlling for other factors. For this exercise, we use the average age of (top) managers (e.g., CEO and CFO) in the 25 largest listed companies in the country (when available), which we collected from publicly available sources. We find a fairly stable relationship between manager age and creativity of innovations at the country level as well, suggesting that the forces we emphasize might account for cross-country differences in the type and quality of innovations.

The cross-country context is also useful for us because it provides a corroboration that manager age is indeed capturing practices related to openness to disruption. We do this by utilizing the individualism and uncertainty avoidance indices of "national cultures" constructed by the Dutch social scientist Geert Hofstede.[3] Our results using these indices are similar to those based on average manager age, suggesting that, at least at the country level, our manager age variable is likely to be capturing some aspects related to a society's openness to disruption.

Our paper is related to several literatures. First, we build on and extend the emerging literature on the interplay between micro and macro aspects of innovation. In particular, we build on Klette and Kortum's (2004) model of innovation dynamics by including a choice between radical and incremental innovations, and by incorporating the dimension of matching between managers of different vintages of human capital (age) and type of innovation.[4] The burgeoning empirical literature in this area (e.g., Foster, Haltiwanger and Krizan, 2001, Lentz and Mortensen, 2008, Akcigit and Kerr, 2010, Hurst and Pugsley, 2011, Syverson, 2011, Kogan et al., 2012, Acemoglu et al., 2013) focuses on R&D, patent and productivity dynamics. We depart from this literature both by focusing on the choice between radical (creative) and incremental innovations, and by presenting a detailed analysis of the relationship between creativity of innovations and manager age.

Second, three papers most closely related to our work are MacDonald and Weisbach (2004), Gorodnichenko and Roland (2012) and Fogli and Veldkamp (2013). MacDonald and Weisbach construct an overlapping generations model in which each generation makes technology-specific human capital investments. They show that younger agents are the ones who invest in human capital complementary to new technologies. Their framework does not incorporate innovations and thus has no distinction between creative, radical innovations vs. incremental innovations. Gorodnichenko and Roland draw a link between innovation and individualism and provide evidence using Hofstede's individualism data. Despite the similar motivating questions, the approaches of the two papers are very different. While Gorodnichenko and Roland look at aggregate measures of productivity, such as TFP or labor productivity, we focus on creative innovations defined using patent citations data from the USPTO. We therefore first start with a microeconomic model of how firms choose their innovation strategies and how managers of different ages endogenously sort across different types of firms. Our main empirical work instead uses the proxy for openness to disruption we have constructed ourselves based on the age of managers across countries and, more centrally, focuses on firm-level and patent-level analysis across US companies. Fogli and Veldkamp also use Hofstede's individualism index in their theoretical and empirical analysis of "individualistic" social

[3] The individualism index is based on Durkheim's (1933) distinction between collectivism and individualism, and measures the extent to which a society functions by relying on loosely knit social ties and thus permits and condones individual actions even when they conflict with collective goals and practices, particularly in a business context. The uncertainty avoidance index, on the other hand, is an inverse proxy for a society's tendency for risk-taking based in part on ideas from Cyert and March's seminal (1963) book.

[4] This matching aspect is common with theoretical analyses of the role of managers, in particular, Lucas (1978), Garicano (2000), and Garicano and Rossi-Hansberg (2004).

networks and the diffusion of new technologies, but their emphasis is on how new technologies diffuse over different network structures and their empirical work exploits exposure to different types of diseases to generate cross-country variation in societal network structures.

Third, our work is linked to the small literature on age and creativity. Galenson and Weinberg (1999, 2001), Weinberg and Galenson (2005), Jones and Weinberg (2011) and Jones (2010) provide evidence that a variety of innovators and top scientists are more creative early in their careers, but they also acquire other types of human capital (perhaps generating different types of creativity) later on. Jones (2009) develops a model in which scientists have to spend more time mastering a given area and have to work in teams because the existing stock of knowledge is growing and thus becoming more difficult to absorb and use. Relatedly, Sarada and Tocoian (2013) investigate the impact of the age of the founders of a company on subsequent performance using Brazilian data.[5]

Fourth, our work is related to the literature pioneered by Bertrand and Schoar (2003) and Bloom and Van Reenen (2007, 2010) which investigates the relationship between CEO and manager characteristics and firm performance. Benmelech and Frydman (2014), for example, show that military CEOs pursue more conservative investment and financial strategies (lower investment in R&D), are less likely to be involved in financial fraud, and perform better during times of distress. Bennedsen et al. (2008) show that the death of a CEO or shocks to the CEO that potentially affect her focus (death of an immediate family member) impact profitability or operating returns. Kaplan et al. (2012) provide evidence from a factor analysis that CEO ability is positively correlated with subsequent firm performance. Also noteworthy in this context is Barker and Mueller (2002), who show that firms with younger CEOs spend more on R&D.

Finally, there is a growing literature on the impact of cultural factors and practices on long-run economic development. The distinction between individualist and collectivist cultures is deep-rooted in sociology (e.g., Durkheim, 1933) and has been widely applied within the sociology, anthropology and psychology literatures (e.g., Parsons, 1949, Kluckhohn and Strodtbeck, 1961, Schwartz, 1994, Triandis, 1995, and Hofstede, 2001). It has been emphasized within the economics literature by Greif (1994), though we are not aware of any other studies emphasizing or empirically investigating the impact of "openness to disruption".[6] More closely related to our focus in this context is Schumpeter's (1934) vision of an innovator as creating disruption, partly in response to economic incentives and partly for psychological motives that lead them to seek challenges and

[5] See also Azoulay, Manso and Zivin (2011) who document the impact of changes in incentives driven by large academic awards and grants on creativity, and Azoulay, Zivin and Wang (2010) who investigate the impact of the death of a very productive co-author on academic productivity.

There is also an extensive literature in social psychology, mostly using survey and experimental evidence, on age and various attitudes both in general and in business. See, e.g., the survey by Walter and Scheibe (2013).

[6] Other aspects of cultural practices have been emphasized as major determinants of economic developments by, among others, Tabellini (2008a,b), Fernandez and Fogli (2009), Guiso, Sapienza and Zingales (2010), and Alesina, Giuliano and Nunn (2011).

deviate from norms, is more closely related to our focus. Traces of this approach can also be seen in Adorno et al.'s (1950) psychological study of authoritarianism, and in McClelland's (1961) and Winslow and Solomon's (1987) approaches to entrepreneurship (see Kirzner, 1997, for a survey). These ideas have been applied in a cross-country context by Shane (1993, 1995), Hofstede (2001), Schwartz (1994), Schwartz and Bilsky (1990) and others. To the best of our knowledge, no other work links these ideas to creative innovations, develops a formal theory along the lines of what we are attempting here, or provides systematic evidence based on firm- or patent-level data.

The rest of the paper is organized as follows. The next section presents our motivating model. Section 3 describes our data sources and variable construction and provides a few basic descriptive statistics. Section 4 presents our main empirical results, which are based on firm-level data. Section 5 returns to the cross-country data and shows that the patterns we identify in the microdata appear to aggregate up to the country level. Section 6 concludes.

2 Motivating Theory

In this section, we provide a simple model of radical and incremental innovations to motivate both the conceptual underpinnings of our approach and some of our empirical strategies.

2.1 Production

We consider a continuous-time economy in which discounted preferences are defined over a unique final good $Y(t)$. This final good is produced by labor and a continuum of intermediate goods j, each located along a circle, $\mathcal{C}$, of circumference 1. The production technology takes the following constant elasticity of substitution form

$$Y(t) = \frac{1}{1-\beta}\left(\int_{\mathcal{C}} q_j(t)^{\beta} k_j(t)^{1-\beta}\, dj\right) L^{\beta}, \tag{1}$$

where $k_j(t)$ denotes the quantity and $q_j(t)$ the quality (productivity) of intermediate good j used in final good production at time t, while L is the total amount of production labor, which is supplied inelastically.

We follow Klette and Kortum (2004) in defining a firm as a collection of leading-edge (best) technologies. A perfectly enforced patent for each leading-edge quality technology is held by a firm, which can produce it at constant marginal cost γ in terms of the unique final good. Because costs and revenues across product lines are independent, a firm will choose price and quantity to maximize profits on each of its product lines. In doing so, it will face an iso-elastic inverse demand derived from equation (1), which can be written, suppressing time arguments, as:

$$p_j = L^{\beta} q_j^{\beta} k_j^{-\beta}, \forall j \in \mathcal{C}.$$

The profit-maximization problem of the firm with leading-edge technology for intermediate good j can then be written as

$$\Pi(q_j) = \max_{k_j \geq 0} \left\{ L^\beta q_j^\beta k_j^{1-\beta} - \gamma k_j \right\} \ \forall j \in \mathcal{C}.$$

The first-order condition of this maximization problem implies a constant markup over marginal cost, $p_j = \gamma/(1-\beta)$, and thus

$$k_j = \left[\frac{(1-\beta)}{\gamma} \right]^{\frac{1}{\beta}} L q_j. \tag{2}$$

Equilibrium profits for a product line with technology q_j are

$$\begin{aligned} \Pi(q_j) &= \beta \left[\frac{(1-\beta)}{\gamma} \right]^{\frac{1-\beta}{\beta}} L q_j \\ &\equiv \pi q_j, \end{aligned}$$

where the second line defines π.

2.2 Managers

In addition to workers, the economy is also populated by managers, who play both an operational role (reducing costs for firms) and manage innovation.

Managers enter and exit the economy following a stationary Poisson birth and death process, so that the measure of managers, M, and their age distribution is constant over time. We index a manager by her age a, or equivalently by her birth date b. Denoting the death rate of managers by δ, the fact that the measure of managers is constant at M implies that the age distribution of managers is simply given by an exponential distribution, i.e., the fraction of managers who are below the age a is $1 - e^{-\delta a}$.[7]

When a manager is born, she acquires the knowledge associated with the average technology in the period in which she is born, giving her a knowledge base of

$$\bar{q}_b \equiv \int_{\mathcal{C}} q_{jb} dj.$$

Similarly, we denote the current period's knowledge stock—current average technology—by

$$\bar{q}_t \equiv \int_{\mathcal{C}} q_{jt} dj.$$

Managers will be hired by monopolists to manage production and innovation in their leading-edge products. In equilibrium, managers will be paid a wage $w_{b,t}$ as a function of the current period's technology, $\bar{q}_t$, and their knowledge, $\bar{q}_b$. We assume that $M < 1$, which implies that the measure of managers is less than the measure of product lines in the economy, so some product lines will

[7] It is also straightforward to see that if the birth rate of managers is given by δ^{birth}, then $M = \delta^{birth}/\delta$.

not use a manager. This simplifies the analysis by providing a simple boundary condition for the determination of equilibrium wages of managers. We also assume that M is not too small, which will ensure that all firms that need a manager for a "radical innovation," as described next, are able to hire one (one can take $M \to 1$ without any loss of generality).

2.3 Corporate Culture and Innovation Dynamics

The economy is populated by two types of firms, with firm type denoted by $\theta \in \{\theta_H, \theta_L\}$ where $\theta_H > \theta_L$. Firm type does not affect productivity directly, but influences the success of radical innovations. In particular, high-type firms, i.e., those with $\theta = \theta_H$, are those with corporate cultures that are open to disruption, and will thus have a comparative advantage in radical innovations. In contrast, we will suppose that low-type firms, i.e., those with $\theta = \theta_L$, is incapable of engaging in radical innovations, thus setting $\theta_L = 0$. Firm type is initially determined upon entry (as described in the next subsection). Thereafter, a low-type firm switches to high type at flow rate $\varphi \in (0, 1)$.[8]

The productivity of each intermediate product is determined by its location along a quality ladder in a given product line. In addition, as noted above and following Klette and Kortum (2004), each leading-edge technology gives the firm an opportunity for further innovation. Innovation dynamics at the firm level are determined by whether the firm pursues an *incremental innovation* or a *radical innovation* strategy. Low-type firms can only engage in incremental innovations as we describe next.

Incremental Innovation Incremental innovations improve the productivity of a product line within the current *technology cluster*.[9] A technology cluster here refers to a specific family of technologies for that product line. Because incremental innovations take place within this technology cluster, they will run into diminishing returns. We model this by assuming that the additional productivity improvements generated by an innovation decline in the number of prior incremental innovations within a technology cluster. In addition, again for the same reason, incremental innovations build on a narrow technology base and create improvements only over this base. This implies that, as illustrated in Example 1 below, incremental innovations will have few citations and limited "generality" (captured by the dispersion of citations they receive from different technology classes as we discuss further below).

We assume that all firms (regardless of their type) can successfully innovate incrementally at the exogenous rate $\xi > 0$. The n^{th} incremental innovation in a technology cluster improves the current productivity of product line j by a step size $\eta_n(q_j, \bar{q}_t)$, where q_j is the current productivity

[8] We assume that there are no switches from high type to low type to simplify the expressions and the analysis.
[9] Our modeling of technology clusters follows Akcigit and Kerr (2010) and Abrams et al (2013).

of the technology, and $\bar{q}_t$ is the current period's technology, and

$$\eta_n(q_j, \bar{q}_t) = [\kappa \bar{q}_t + (1-\kappa) q_j] \eta \alpha^n \tag{3}$$

with $\alpha \in (0,1)$, $\eta > 0$, and $\kappa \in (0,1)$. This functional form implies two features. First, each innovation builds both on the current productivity of the product line where it originates, with weight $1-\kappa$, and on average technology, $\bar{q}_t$, with weight κ. Second, productivity gains from incremental innovations decline geometrically, at the rate α, in the number of prior incremental innovations in the technology cluster.

Denoting by t_n the time of the n^{th} incremental improvement for product line j, the evolution of the technology of product line j in a technology cluster that started with productivity q_j^0 after n incremental innovations can then be written as

$$\begin{aligned} q_j^n &= q_j^0 + \sum\nolimits_{i=0}^{n} \left[\kappa \bar{q}(t_i) + (1-\kappa) q_j^0\right] \eta \alpha^i \\ &= q_j^0 \left[1 + (1-\kappa)\eta \frac{1-\alpha^{n+1}}{1-\alpha}\right] + \eta\kappa \sum\nolimits_{i=0}^{n} \alpha^i \bar{q}(t_i). \end{aligned}$$

Radical Innovations Radical innovations combine the current technology of the product line the firm is operating, the knowledge base of the manager, and the available knowledge stock of the economy to innovate in a new area (creatively destroying the leading-edge technology of some other firm). Similar to Weitzman's (1998) approach based on recombination, this combination of knowledge bases creates a new technology cluster. Because they create new technology clusters, radical innovations tend to receive more citations, are more likely to have a very high number of ("tail") citations, and have greater generality.

If there is a radical innovation in a particular product line, the innovator will initiate a new technology cluster in a different product line (and will still keep its original product line). The creation of a new technology cluster generates a larger improvement on current technology, and also provides the innovator with the opportunity to start a new series of incremental innovations. Because radical innovations are not directed and each firm controls an infinitesimal fraction of all products, the likelihood that it will be the firm itself radically innovating over its own product is zero.[10] Thus radical innovations are associated with "Schumpeterian creative destruction." We next describe the technology for radical innovations.

A successful radical innovation leads to an improvement over the product line uniformly located on the circle $\mathcal{C}$, and thus generates creative destruction. In particular, if there is a successful radical innovation over a product line with technology q_j, this leads to the creation of a new leading-edge

[10] It may be more plausible to assume that radical innovations also take place over a range of products that are "technologically close" to the knowledge base of the innovator. Provided that there is a continuum of products within this range, this would not affect any of our results.

technology (now under the control of the innovating firm and manager), with productivity

$$q_j^0 = q_j + \eta_0,$$

where the superscript 0 denotes the fact that a radical innovation initiates a new cluster with no prior incremental innovations.

Managers' Role For each of their active product lines, firms hire managers who influence their revenues in two ways. First, a manager of age $a = t - b$ contributes $\bar{q}_t f(a)$ to the revenues of a firm when the aggregate technology level is $\bar{q}_t$ (e.g., by reducing costs).[11] We presume (but do not need to impose) that f is increasing, so that more experienced managers are better at cost reductions. If the firm hires no manager, then it does not receive this additional revenue. Second, a manager affects the flow rate of radical innovations for firms attempting such radical innovations, as we describe next.

A firm of type θ has a baseline flow rate of radical innovation (regardless of whether they are pursuing radical or incremental innovations) equal to $\psi\Lambda\theta$. In addition, if it pursues a radical innovation strategy, hires a manager with knowledge $\bar{q}_b$ and the current technology in the economy is $\bar{q}_t$, it will also have a flow rate of radical innovation equal to

$$\Lambda\theta\bar{q}^a, \tag{4}$$

where

$$\bar{q}^a \equiv \frac{\bar{q}_b}{\bar{q}_t}$$

is the relative average quality of managers of age a, and $\Lambda \in (0,1]$ (and the superscript, rather than a subscript, here emphasizes that this is a ratio of two averages). This specification implies that low-type firms, with $\theta_L = 0$, cannot engage in radical innovations—i.e., both $\psi\Lambda\theta_L$ and $\Lambda\theta_L$ are equal to zero.

Moreover, since both high- and low-type firms have the same rate of success, at the rate ξ, when they attempt incremental innovations, our model also implies that θ captures the *comparative advantage* of firms for radical innovation. In addition, young managers also have a *comparative advantage* in radical innovation—since the contribution of the manager of age a to cost reductions is the same for all firms, and younger managers contribute to the flow rate of radical innovation with high-type firms.

The parameter Λ captures the role of institutional or social sanctions on radical innovations. Such sanctions may permit only the implementation of certain radical innovations, thus making successful innovations less likely.[12]

[11] We model this contribution as an additive element in the revenues of the firm so as not to affect its monopoly price and quantity choices of the firm via this channel.

[12] In particular, in the context of our modeling of product lines along the circle $\mathcal{C}$, we may assume that such sanctions

Radical Innovations and Citation Patterns The next example provides more details on the evolution of technology clusters and the citation pattern for the patents related to the incremental and radical innovations therein.

Example 1 The following chart provides an illustrative example focusing on two product lines:

$$\text{First product line:} \quad \Big| \; \underbrace{\overbrace{P_1}^{\eta_0}\ \overbrace{P_2}^{\eta_1}\ \overbrace{P_6}^{\eta_2}\ \overbrace{P_7}^{\eta_3}}_{\textit{Tech Cluster 1}} \; \Big| \; \underbrace{\overbrace{P_{11}}^{\eta_0}\ \overbrace{P_{12}}^{\eta_1}}_{\textit{Tech Cluster 2}} \; \Big| \; \underbrace{\overbrace{P_{13}}^{\eta_0}\ \overbrace{P_{14}}^{\eta_1}\ \overbrace{P_{15}}^{\eta_2}}_{\textit{Tech Cluster 3}}$$

$$\text{Second product line:} \quad \Big| \; \underbrace{\overbrace{P_3}^{\eta_0}\ \overbrace{P_4}^{\eta_1}\ \overbrace{P_5}^{\eta_2}}_{\textit{Tech Cluster 1}} \; \Big| \; \underbrace{\overbrace{P_8}^{\eta_0}\ \overbrace{P_9}^{\eta_1}\ \overbrace{P_{10}}^{\eta_2}}_{\textit{Tech Cluster 2}}$$

In this example, P_n denotes the n^{th} patent registered at the patent office and η_n denotes the step size as described in equation (3). The first technology cluster starts with a radical innovation associated with a patent P_1. The productivity improvement due to this patent is η_0. Subsequently a new incremental innovation in this technology cluster, with patent P_2, follows on P_1, increasing productivity by another $\eta_1 < \eta_0$. After this innovation, there is a radical innovation P_3 in the second product line, followed by two subsequent incremental innovations P_4 and P_5. Since P_5 and P_6 are second incremental innovations in their technology clusters, they increase productivity by $\eta_2 < \eta_1$. Note that P_1, P_3, P_8, P_{11} and P_{13} are radical innovations starting new technology clusters. As described above, these come from innovations in other product lines operated by high-type firms. Suppose also that the firm operating technology cluster 1 with patent P_7 is a high-type firm, and successfully undertakes a radical innovation after P_7, launching a new technology cluster on a different product line (shown above as patent P_8*).*

Consider next the patterns of citation resulting from these innovations. It is natural to assume that each incremental innovation will cite all previous innovations in its technology cluster, which is the pattern shown in the next table. (Alternatively, such patents might also cite patterns from previous technology clusters on the same product line, with very similar patterns). In addition, it is also plausible that, because a radical innovation is recombining ideas from its own product line and the product line on which it is building, it should be citing the fundamental ideas encapsulated in the patents that initiated the two technology clusters. For this reason, patents P_8, P_{11}, and P_{13} cite the patents initiating the previous technology cluster in this product line as well as the patent

permit a firm operating product line j to successfully innovate over technologies that are sufficiently close to itself. Suppose, for example, that j may be allowed to innovate only on product lines that are at most a distance Λ from itself. Then the case of no restrictions would correspond to $\Lambda = 1/2$, so that radical innovations over any product lines on the circle $\mathcal{C}$ are possible, while $\Lambda < 1/2$ would correspond to restrictions and thus lower the likelihood of successful radical innovations.

initiating the most recent technology cluster in their own product line. The next table shows this citation pattern for the first five patents.

Cited	***Citing***
P_1 :	$P_2, P_6, P_7, P_8, P_{11}$
P_2 :	P_6, P_7
P_3 :	P_4, P_5, P_8
P_4 :	P_5
P_5 :	*none*

For example, P_2 builds only on P_1 and thus only cites P_1, and is in turn cited by P_6 and P_7. P_1 is cited not only by the patents that build on itself within the same product line, P_2, P_6, P_7 and P_{11}, but also by P_8 because this new patent comes out of recombining ideas based on this technology cluster and those in some other product line. This pattern then implies that radical innovations will receive more citations and will receive more "general" citations as well. They will also be heavily overrepresented among "tail innovations," meaning among patents receiving the highest number of citations. These are the patterns we will explore in our empirical work.

We close the model by assuming that new firms enter at the exogenous flow rate $x > 0$, and entry corresponds to a (radical) innovation over an existing product line uniformly at random, which thus initiates a new technology cluster. We further assume that immediately after entry, a firm's type is also drawn at random. In particular, successful entrant is high-type, $\theta = \theta_H$, with probability $\zeta \in (0, 1)$, and is low-type, $\theta = \theta_L$ $(= 0)$, with the complementary probability, $1 - \zeta$. Thereafter, firm type (corporate culture) evolves according to the Markov chain described above.

2.4 Value Functions and Firm Maximization

Recall that a firm makes the innovation decision in each of its product lines to maximize its present discounted value, which we denote by $W_s(\overrightarrow{q_f}, \overrightarrow{n_f})$ where $s \in \{H, L\}$, $\overrightarrow{q_f}$ is the vector of productivities of the firm, $\overrightarrow{n_f}$ is the vector of the number of incremental innovations in each of these product lines, i.e., $\overrightarrow{q_f} \equiv \left\{q_{f,j_1}, q_{f,j_2}, ..., q_{f,j_{m_f}}\right\}$, and $\overrightarrow{n_f} \equiv \left\{n_{f,j_1}, n_{f,j_2}, ..., n_{f,j_{m_f}}\right\}$, and m_f denotes the number of product lines that firm f is operating.[13] The value function for a low-type firm can be written as

$$rW_L(\overrightarrow{q_f}, \overrightarrow{n_f}) - \dot{W}_L(\overrightarrow{q_f}, \overrightarrow{n_f}) = \sum_{m=1}^{m_f} \left[\begin{array}{c} \max_a \{\pi q_{f,j_m} + \bar{q}_t f(a) - w_{a,t}\} \\ +\xi \left[\begin{array}{c} W_L \left(\begin{array}{c} \overrightarrow{q_f} \setminus \{q_{f,j_m}\} \cup \left\{q_{f,j_m} + \eta_{n_{f,j_m}+1}\right\}, \\ \overrightarrow{n_f} \setminus \{n_{f,j_m}\} \cup \{n_{f,j_m} + 1\} \end{array} \right) \\ -W_L(\overrightarrow{q_f}, \overrightarrow{n_f}) \end{array} \right] \\ +\tau \left[W_L\left(\overrightarrow{q_f} \setminus \{q_{f,j_m}\}, \overrightarrow{n_f} \setminus \{n_{f,j_m}\}\right) - W_L(\overrightarrow{q_f}, \overrightarrow{n_f})\right] \end{array} \right] + \varphi \left[W_H(\overrightarrow{q_f}, \overrightarrow{n_f}) - W_L(\overrightarrow{q_f}, \overrightarrow{n_f})\right]. \quad (5)$$

[13] Here and elsewhere, we suppress time as an explicit argument of the value functions to simplify notation.

We can explain the right-hand side of this value function as follows: for each product line $m = 1, ..., m_f$, the firm receives a revenue stream of $\pi q_{f,j_m}$ as a function of its productivity in this product line, q_{f,j_m}. In addition, it has a choice of the age of the manager it will hire to operate this product line (formally choosing $a \in \mathbb{R}_+ \cup \{\emptyset\}$, which is suppressed to save on notation), and if the manager's age is a, it will have additional revenue/cost savings of $\bar{q}_t f(a)$ and pay the market price for such a manager of age a at time t, $w_{a,t}$. Summing over all of its product lines gives the current revenues of the firm. In addition, the firm can undertake an innovation on the basis of each of its active product lines. Since we are looking at a low-type firm, all innovations will be incremental, thus arriving at the rate ξ. When such an innovation happens in product line m that has already undergone n_{f,j_m} incremental innovations, the mth element of $\overrightarrow{q_f}$ changes from q_{f,j_m} to $q_{f,j_m} + \eta_{n_{f,j_m}+1}$ and n goes up by one. We represent this with the arguments of the value function changing to $\overrightarrow{q_f} \backslash \{q_{f,j_m}\} \cup \left\{q_{f,j_m} + \eta_{n_{f,j_m}+1}\right\}, \overrightarrow{n_f} \backslash \{n_{f,j_m}\} \cup \{n_{f,j_m} + 1\}$ (and the firm relinquishes its current value function $W_L(\overrightarrow{q_f}, \overrightarrow{n_f})$). The firm might also lose one of its currently active product lines to creative destruction, which happens at the endogenous rate τ (which will be determined in Section 2.6), and in that case, the firm's value function changes from $W_L(\overrightarrow{q_f}, \overrightarrow{n_f})$ to $W_L(\overrightarrow{q_f} \backslash \{q_{f,j_m}\}, \overrightarrow{n_f} \backslash \{n_{f,j_m}\})$ (i.e., $\overrightarrow{q_f}$ changes $\overrightarrow{q_f} \backslash \{q_{f,j_m}\}$ and $\overrightarrow{n_f}$ to $\overrightarrow{n_f} \backslash \{n_{f,j_m}\}$). Finally, the last term is due to the fact that a low-type firm switches to high-type at the flow rate φ, in which case it relinquishes its current value function and begets the value function of a high-type firm, $W_H(\overrightarrow{q_f}, \overrightarrow{n_f})$.

The value function of a high-type firm can be similarly written as

$$
\begin{aligned}
& rW_H(\overrightarrow{q_f}, \overrightarrow{n_f}) - \dot{W}_H(\overrightarrow{q_f}, \overrightarrow{n_f}) \qquad (6) \\
&= \sum_{m=1}^{m_f} \max \left\{ \begin{array}{c} \pi q_{f,j_m} \\ + \max_a \left\{ \bar{q}_t f(a) - w_{a,t} + \xi \left[\begin{array}{c} W_H \left(\begin{array}{c} \overrightarrow{q_f} \backslash \{q_{f,j_m}\} \cup \left\{ q_{f,j_m} + \eta_{n_{f,j_m}+1} \right\}, \\ \overrightarrow{n_f} \backslash \{n_{f,j_m}\} \cup \{n_{f,j_m}+1\} \end{array} \right) \\ -W_H(\overrightarrow{q_f}, \overrightarrow{n_f}) \end{array} \right] \right\}; \\ \pi q_m + \max_a \left\{ \bar{q}_t f(a) + \Lambda \theta_H \bar{q}^a \left[\begin{array}{c} \mathbb{E} W_H \left(\overrightarrow{q}_f \cup \{q_{j'} + \eta_0\}, \overrightarrow{n}_f \cup \{0\} \right) \\ -W_H(\overrightarrow{q}_f, \overrightarrow{n}_f) \end{array} \right] - w_{a,t} \right\} \end{array} \right\} \\
&+ \sum_{m=1}^{m_f} \tau \left[W_H(\overrightarrow{q_f} \backslash \{q_{f,j_m}\}, \overrightarrow{n_f} \backslash \{n_{f,j_m}\}) - W_H(\overrightarrow{q_f}, \overrightarrow{n_f}) \right] \\
&+ \sum_{m=1}^{m_f} \psi \Lambda \theta_H \left[\mathbb{E} W_H \left(\begin{array}{c} \overrightarrow{q}_f \cup \{q_{j'} + \eta_0\}, \\ \overrightarrow{n}_f \cup \{0\} \end{array} \right) - W_H(\overrightarrow{q}_f, \overrightarrow{n}_f) \right]
\end{aligned}
$$

The intuition for this value function is very similar to (5) except for the possibility of a radical innovation. In particular, for each product line m, this high-type firm has a radical innovation at the flow rate $\psi \Lambda \theta_H$ regardless of its innovation strategy. In addition it has a choice between incremental and radical innovation, represented by the outer maximization. The first option here is choosing incremental innovation for product line m and is thus similar to the first line of (5). The second option is radical innovation, and in this case the trade-off involved in the age of the manager is different, since manager age affects the arrival rate of radical innovations as shown in (4). In the case of a successful radical innovation, the value of the firm changes to $\mathbb{E}W_H\left(\overrightarrow{q_f} \cup \left\{q_{j'} + \eta_0\right\}, \overrightarrow{n_f} \cup \{0\}\right)$, where the expectation is over a product line drawn uniformly at random upon which the radical innovation will build.

The next proposition shows that, as in Klette and Kortum (2004) and Acemoglu et al. (2013), these value functions can be decomposed into sums of value functions defined at the product-line level.

Proposition 1 *The value functions in* (5) *and* (6) *can be written as*

$$W_s\left(\overrightarrow{q_f}, \overrightarrow{n_f}\right) = \sum_{m=1}^{m_f} V_s\left(q_j, n\right),$$

where $V_s(q_j, n)$ *is the (franchise) value of a product line of productivity* q_j *with* n *incremental innovations that belongs to a firm of type* $s \in \{H, L\}$ *such that*

$$\begin{aligned} rV_L\left(q_j, n\right) - \dot{V}_L\left(q_j, n\right) &= \max\left\{\pi q_j + \bar{q}_t f(a) - w_{a,t}\right\} + \xi\left[V_L\left(q_j + \eta_{n+1}, n+1\right) - V_L\left(q_j, n\right)\right] \\ &\quad - \tau V_L\left(q_j, n\right) + \varphi\left[V_H\left(q_j, n\right) - V_L\left(q_j, n\right)\right], \end{aligned} \tag{7}$$

and

$$\begin{aligned} & rV_H\left(q_j, n\right) - \dot{V}_H\left(q_j, n\right) \\ = \; & \max\left\{ \begin{array}{c} \pi q_j + \max_a \left\{ \bar{q}_t f(a) - w_{a,t} + \xi \left[\begin{array}{c} V_H\left(q_j + \eta_{n+1}, n+1\right) \\ -V_H\left(q_j, n\right) \end{array} \right] \right\}; \\ \pi q_j + \max_a \left\{ \bar{q}_t f(a) + \Lambda \theta_H \bar{q}^a \mathbb{E} V_H\left(\bar{q}_t\right) - w_{a,t} \right\} \end{array} \right\} \\ & - \tau V_H\left(q_j, n\right) + \psi \Lambda \theta_H \mathbb{E} V_H\left(\bar{q}_t\right), \end{aligned} \tag{8}$$

where $\mathbb{E}V_H(\bar{q}_t)$ *denotes the expected value of a radical innovation when the aggregate technology level is* $\bar{q}_t$*.*

Proof. Both of these value functions can be derived straightforwardly by conjecturing the above forms and verifying the conjecture. ■

2.5 Stationary Equilibrium With $\kappa = 1$

We now characterize the stationary equilibrium of this economy in the case where $\kappa = 1$—so that all current innovations build on current technology, $\bar{q}_t$ (and not on the current productivity of the existing technology cluster). This assumption considerably simplifies the analysis, and we return to the general case where $\kappa < 1$ below.

Value Functions in Stationary Equilibrium A *stationary equilibrium* is defined as an equilibrium in which aggregate output, Y_t, grows at a constant rate g, and the distribution of product lines between high- and low-type firms and over the prior number of incremental innovations remains stationary.

As noted above, firms decide the age of the manager to hire for each of the product lines they are operating and whether to engage in a radical or incremental innovation. Let us first consider the value of a product line for a low-type firm. From Proposition 1, we can focus on the decisions and the value function of such a firm at the product line level, and the relevant value function is given by (7).

Since some firms will not hire managers (as $M < 1$), all firms not undertaking radical innovations must be indifferent between hiring and not hiring a manager, which implies that the equilibrium wage for managers, employed by firms engaged in incremental innovations, satisfies the boundary condition:

$$w_{a,t} = \bar{q}_t f(a). \tag{9}$$

Substituting the equilibrium wage (9) into (7), we obtain a simplified value function for low-type firms as

$$\begin{aligned} rV_L(q_j, n) - \dot{V}_L(q_j, n) &= \pi q_j + \xi\left[V_L\left(q_j + \bar{q}_t \eta \alpha^{n+1}, n+1\right) - V_L(q_j, n)\right] \\ &\quad -\tau V_L(q_j, n) + \varphi\left[V_H(q_j, n) - V_L(q_j, n)\right]. \end{aligned}$$

Solving this value function gives an explicit characterization of the value function of low-type firms as shown in the next proposition.

Proposition 2 *Let us assume that the value function for a high-type firm takes the following form:* $V_H(q_j, n) = Aq_j + \tilde{B}(n)\bar{q}_t$. *Then the value function of a product line operated by a low-type firm,* (7) *takes the following form*

$$V_L(q_j, n) = Aq_j + B(n)\bar{q}_t \tag{10}$$

where

$$A \equiv \frac{\pi}{r+\tau};\ [r - g + \xi + \tau + \varphi] B(n) = \xi A \eta \alpha^{n+1} + \varphi \tilde{B}(n) + \xi B(n+1);$$

and $\tilde{B}(n)$ *is defined in Proposition 3 below.*

Proof. See the Appendix. ■

The form of the value function in (10) is intuitive. It depends linearly on current productivity, q_j, since this determines the current flow of profits. It also depends on current economy-wide technology, $\bar{q}_t$, since all innovations, including incremental ones, build on this. Finally, it is decreasing in n (because $B(n)$ is decreasing) since a higher n implies that the next incremental innovation will increase productivity by less—and incremental innovation is the only type of innovation that a low-type firm can undertake.

We next turn to the value of a product line operated by a high-type firm, which differs from (7) because high-type firms have to decide whether to engage in incremental or radical innovation, given by (8) above. Because (4) implies that younger managers have comparative advantage in radical innovation, it follows straightforwardly that there will exist a maximum age a^* such that only managers below this age will work in firms attempting radical innovation. Moreover, the maximization over the age of the manager in (8) implies that such a firm must be indifferent between hiring any manager younger than a^*. This implies:

$$\bar{q}_t f(a^*) + \Lambda\theta_H \bar{q}^{a^*}\mathbb{E}V_H(\bar{q}_t) - w_{a^*,t} = \bar{q}_t f(a) + \Lambda\theta_H \bar{q}^{a}\mathbb{E}V_H(\bar{q}_t) - w_{a,t} \text{ for all } a < a^*.$$

Note that the oldest manager hired for radical innovation earns (from expression (9))

$$w_{a^*,t} = \bar{q}_t f(a^*).$$

Hence

$$w_{a,t} = \begin{cases} \bar{q}_t f(a) & \text{for} \quad a > a^* \\ \bar{q}_t f(a) + \Lambda\theta_H \left[\bar{q}^a - \bar{q}^{a*}\right]\mathbb{E}V_H(\bar{q}_t) & \text{for} \quad a \leq a^* \end{cases}. \tag{11}$$

This wage schedule highlights that, in general, younger or older managers might be paid more (this will depend on the f function). Younger managers have a comparative advantage in radical innovation, but older managers might be more productive in operating firms.[14]

Now substituting for (11) in (8), we obtain a simplified form of the value function of a product line operated by a high-type firm as

$$\begin{aligned} rV_H(q_j,n) - \dot{V}_H(q_j,n) \quad = \quad & \max\left\{ \begin{array}{c} \pi q_j + \xi\left[V_H\left(q_j + \bar{q}_t\eta\alpha^{n+1}, n+1\right) - V_H(q_j,n)\right]; \\ \pi q_j + \Lambda\theta_H \bar{q}^{a^*}\mathbb{E}V_H(\bar{q}_t) \end{array} \right\} \\ & -\tau V_H(q_j,n) + \psi\Lambda\theta_H\mathbb{E}V_H(\bar{q}_t). \end{aligned}$$

We next characterize the solution to this value function and also determine the allocation of managers to different product lines (and to incremental and radical innovations).

Proposition 3 *The value function in* (8) *takes the following form*

$$V_H(q_j,n) = Aq_j + \bar{q}_t\tilde{B}(n), \tag{12}$$

[14] The evidence in Galenson and Weinberg (1999, 2001), Weinberg and Galenson (2005) and Jones and Weinberg (2011) is consistent with the possibility that either younger or older creative workers might be more productive.

where A and $B(n)$ are as defined in Proposition 2) and $\tilde{B}(n)$ is given by

$$[r-g+\tau]\tilde{B}(n) = \psi\left[A(1+\eta)+\tilde{B}(0)\right] + \begin{cases} \xi\left[\tilde{A}\eta\alpha^{n+1}+\tilde{B}(n+1)-\tilde{B}(n)\right] & \text{for } n<n^* \\ \Lambda\theta_H\bar{q}^{a^*}\left[(1+\eta)\tilde{A}+\tilde{B}(0)\right] & \text{for } n\geq n^* \end{cases}, \tag{13}$$

where $n^* \in \mathbb{Z}_{++}$ is the number of incremental innovations within a technology cluster at which there is a switch to radical innovation given by

$$n^* = \lceil n' \rceil \text{ such that } \xi\left[A\eta\alpha^{n'+1}+\tilde{B}(n'+1)-\tilde{B}(n')\right] = \Lambda\theta_H\bar{q}^{a^*}\left[(1+\eta)A+\tilde{B}(0)\right]. \tag{14}$$

Proof. See the Appendix. ■

The intuition for this high-type value function is similar to that for Proposition 2, except that the dependence on the number of prior innovations in the current technology cluster, n, is more complicated since when n exceeds n^*, a high-type firm will switch to radical innovation (and from that point on n will no longer be relevant). This critical value n^* is given by (14), which designates the smallest integer after n' where n' equates the value of attempting an additional incremental innovation to the value of attempting a radical innovation (the notation $\lceil n \rceil$ denotes the next integer after n).

It is also worth noting that this threshold, n^*, is constant in the stationary equilibrium. This is because the value function increases linearly in $\bar{q}_t$, but the knowledge stock and wages of managers also increase linearly, and in the stationary equilibrium, these two forces balance out, ensuring that n^* is constant while V_H increases linearly in $\bar{q}_t$.

Given the form of V_H, $\mathbb{E}V_H(\bar{q}_t)$, the value of a new radical innovation, can be written as

$$\begin{aligned} \mathbb{E}V_H(\bar{q}_t) &= \mathbb{E}\left[\tilde{A}q_j+\tilde{A}\eta\bar{q}_t+\bar{q}_t\tilde{B}(0)\right] \\ &= [\tilde{A}(1+\eta)+\tilde{B}(0)]\bar{q}_t \\ &\equiv v\bar{q}_t, \end{aligned}$$

where the last line defines v. Then the equilibrium wage schedule simplifies to:

$$w_{a,t} = \begin{cases} f(a)\,\bar{q}_t & \text{for } a>a^* \\ [f(a)+\Lambda\theta_H(\bar{q}^a-\bar{q}^{a^*})v]\bar{q}_t & \text{for } a\leq a^* \end{cases}. \tag{15}$$

and is thus also linear in $\bar{q}_t$.

Equilibrium Characterization The next proposition provides the characterization of the stationary equilibrium.

Proposition 4 *Low-type firms (those with $\theta = \theta_L$) always hire "old" managers (those with $a > a^*$ or $b < b_t^*$), pursue incremental innovations and never generate radical innovations.*

High-type firms (those with $\theta = \theta_H$) pursue incremental innovations on product lines less than n^* prior incremental innovations, where n^* is given by (14), and hire "old" managers (those with $a > a^*$ or $b < b_t^*$). They pursue radical innovations on product lines with more than n^* and hire "young" managers (those with $a \leq a^*$ or $b \geq b_t^*$).

A lower Λ (corresponding to the society being less permissive to radical innovations) will increase n^* (so that a lower fraction of high-type firms will pursue radical innovation), and will reduce the wages of young managers (because there is less demand for the knowledge of young managers).

Proof. This result directly follows from Propositions 2 and 3. ■

Empirical Implications Our empirical work is inspired by Proposition 4. As explained above, radical innovations will be associated with greater indices of our measures of creative innovations (innovation quality, tail innovations, superstar fraction, and generality). We will first investigate the cross-sectional relationship between manager (CEO) age and creative innovations. In these cross-sectional regressions, manager age is taken to be a proxy of a corporate culture that is more open to disruption. Therefore, from Proposition 4, we expect a negative cross-sectional relationship between manager age and creative innovations. As just stressed, this cross-sectional relationship does not correspond to the "causal effect" of manager age on creativity of innovations (which would apply if we varied manager age holding the firm's corporate culture constant); in particular, it also reflects the sorting of younger managers to corporate cultures that are open to disruption (and thus more conducive to creative innovations).

Our model also has longitudinal implications—that is, implications about how manager age and creativity of innovations vary over time at the firm level—which shed further light on the relative magnitudes of the sorting and the causal effects. To understand these implications, let us consider the innovation dynamics of firms implied by Proposition 4.

Recall that low-type firms always engage in incremental innovations and never generate radical innovations. In contrast, high-type firms may attempt a radical innovation depending on how many prior incremental innovations they have had on a product line.

- For product line with $n < n^*$, a high-type firm hires an old manager (or keeps its already existing old manager), and pursues an incremental innovation strategy. Given the technology specified above, however, such a firm still generates radical innovations at the rate $\psi\Lambda\theta_H$.

- For a product line with $n \geq n^*$, a high-type firm hires a young manager and engages in radical innovation. In this case, the average rate of radical innovation across product lines with $n \geq n^*$ and operated by high-type firms can be computed using the aforementioned fact that the age distribution of managers is given by the exponential distribution, as

$$\psi \Lambda \theta_H + \frac{1}{F(a^*)} \int_0^{a^*} \Lambda \theta_H \bar{q}^a dF(a) = \psi \Lambda \theta_H + \frac{\Lambda \theta_H \delta}{g+\delta} \frac{\left[1 - e^{-(g+\delta)a^*}\right]}{\left[1 - e^{-\delta a^*}\right]}. \tag{16}$$

Now consider a low-type firm that switches to high-type, and to simplify the discussion, suppose that it has a unique product line. Then, if this product line has had $n < n^*$ incremental innovations, the firm will continue to pursue an incremental innovation strategy, keeping its old manager.[15] In the process, it will generate radical/creative innovations at the flow rate $\psi \Lambda \theta_H$ as noted above. When it reaches $n = n^*$, it will hire a young manager, switch to a radical innovation strategy, and at that point, its rate of radical/creative innovations will increase, on average, from $\psi \Lambda \theta_H$ to the expression in (16). In contrast, if the product line of the firm at the time of switching to high-type has had $n \geq n^*$ incremental innovations, it will immediately hire a young manager, pursue a radical innovation strategy, and have radical innovations at the flow rate (16).

This discussion implies that when we focus on the relationship between within-firm changes in manager age and creative innovations, we expect to find two regularities. First, when a firm switches from an older to a younger manager, this should be associated with an increase in creative innovations. Second, firms that switch from an older to a younger manager should, on average, experience an increase in creative innovations even *before* the switch. Namely, before the actual switch to a younger manager, the increase in creative innovations approximately corresponds to $\psi \Lambda \theta_H$, whereas following the switch to a younger manager, there will be a further increase in creative innovations corresponding to the second term in (16). Note, however, that even this further increase following the switch to a younger manager does not correspond to the causal effect of manager age on creative innovations for several reasons; first, for firms with $n \geq n^*$, both events will be taking place at the same time; second, even for firms with $n < n^*$, the impact following the switch to a younger manager still contains the sorting effect and also depends on the matching patterns between managers and firms as indicated by the presence of the terms representing the age distribution of managers; and third, strictly speaking to obtain the causal effect, we need to keep the firm type and number of prior incremental innovations constant, and just change (exogenously) manager age—and this is the exercise will will perform in Section 4.4 below.

Finally, though we will not be able to investigate this directly in our empirical work, the implications of changes in Λ are interesting. A lower value of this parameter naturally reduces radical innovations and, at the same time, decreases the wages of young managers, thus making it look like the society is discriminating against the young; but in fact this is a consequence of the

[15] Strictly speaking this is true under an infinitesimal cost of replacing managers. Otherwise, it could fire its old manager and hire another old manager, with no impact on our results or discussion here.

society discouraging radical innovations.

2.6 General Equilibrium and the Stationary Distribution of Products

We next characterize the stationary distribution of product lines in this economy in terms of the types of their owners and also in terms of the prior number of incremental innovations, and then use these distributions to determine the aggregate growth rate of the economy in the stationary equilibrium.

Let us next denote the fraction of product lines occupied by s-type firms (for $s \in \{L, H\}$) with n prior incremental innovations by μ_n^s (these are not functions of time as we are focusing on a stationary equilibrium). Let us also denote the total creative destruction from s-type firms by τ^s. The stationary distribution of product lines is determined by standard flow equations equating inflows and outflows from each state. For high types, these take the form

$$\begin{array}{rcll} \underline{\textsc{Outflow}} & & \underline{\textsc{Inflow}} & \\ \left(\tau^L + \xi\right)\mu_0^H & = & \tau^H\left(1 - \mu_0^H\right) + \varphi\mu_0^L & \text{for} \quad n = 0 \\ \left(\tau^L + \tau^H + \xi\right)\mu_n^H & = & \xi\mu_{n-1}^H + \varphi\mu_n^L & \text{for} \quad n^* > n > 0 \\ \left(\tau^L + \tau^H\right)\mu_n^H & = & \xi\mu_{n-1}^H + \varphi\mu_n^L & \text{for} \quad n = n^* \\ \left(\tau^L + \tau^H\right)\mu_n^H & = & \varphi\mu_n^L & \text{for} \quad n > n^* \end{array}.$$

Consider the first line corresponding to $n = 0$. Outflows from this state, products with $n = 0$ operated by high-type firms, come from two sources. First, there is creative destruction coming from low-type firms, which takes place at the rate τ^L per product line (and hence multiplied by μ_0^H). Second, the high-type firm operating this product line has a successful incremental innovation, which takes place at the rate ξ (similarly multiplied by μ_0^H). Inflows into this state are due to creative destruction coming from the high-type firm, which takes place at the rate τ^H (multiplied by the fraction of all product lines except those that are already in this state, thus $\left(1 - \mu_0^H\right)$), or due to a low-type firm operating a product line at $n = 0$ changing its type to high, which adds the flow rate $\varphi\mu_0^L$. The other lines are explained similarly, except that creative destruction coming from high-type firms also generates outflows for $n > 1$, and there is no inflow coming from incremental innovations for product lines with $n > n^*$ since high-type firms switch to radical innovation at $n = n^*$.[16]

The flow equations for the low-type product lines can be written similarly and have a similar intuition

$$\begin{array}{rcll} \underline{\textsc{Outflow}} & & \underline{\textsc{Inflow}} & \\ \left(\tau^H + \xi + \varphi\right)\mu_0^L & = & (1 - \mu_0^L)\tau^L & \text{for} \quad n = 0 \\ \left(\tau^L + \tau^H + \xi + \varphi\right)\mu_n^L & = & \xi\mu_{n-1}^L & \text{for} \quad n > 0 \end{array}.$$

[16] These equations are written under the assumption that $n^* > 0$. When $n^* = 0$, high-type firms never undertake incremental innovations, and thus the flow equations become $\tau^L\mu_0^H = (1 - \mu_0^H)\tau^H + \varphi\mu_0^L$ for $n = n^* = 0$, and $\left(\tau^L + \tau^H\right)\mu_n^H = \varphi\mu_n^L$ for $n > n^* = 0$.

The creative destruction rates from low-type and high-type firms, in turn, can be computed as

$$\tau^L = x(1-\zeta) \quad \text{and} \quad \tau^H = x\zeta + M\int_0^{a^*} \Lambda\theta_H \bar{q}^a dF(a) + \psi\Lambda\theta_H \sum\nolimits_{n=0}^{\infty} \mu_n^H,$$

where x is the entry rate, $F(a)$ denotes the stationary distribution of manager age, a^* is the threshold below which managers are hired by firms to perform radical innovations, and $\psi\Lambda\theta_H \sum_{n=0}^{\infty} \mu_n^H$ is the rate of radical innovations for high-type firms which applies regardless of whether they pursue a radical innovation strategy. Low-type firms, on the other hand, generate creative destruction only when they initially enter the economy (since they do not engage in radical innovation). Given these quantities, the total creative destruction rate of the economy is given as

$$\tau = \tau^L + \tau^H.$$

To derive the aggregate growth rate, we combine (1) with (2) to obtain

$$Y = \frac{L}{1-\beta}\left[\frac{(1-\beta)}{\gamma}\right]^{\frac{1-\beta}{\beta}} \bar{q}.$$

The growth rate of the economy is then equal to the growth of the average quality $\bar{q}_t$. After a time interval $\Delta t > 0$, the average quality is given by

$$\bar{q}_{t+\Delta t} = \bar{q}_t + \eta\bar{q}_t\tau\Delta t + \bar{q}_t\eta\xi\Delta t\left[\sum\nolimits_{n=0}^{n^*-1} \mu_n^H \alpha^{n+1} + \sum\nolimits_{n=0}^{\infty} \mu_n^L \alpha^{n+1}\right] + o(\Delta t),$$

where we have used the fact that all radical innovations come from creative destruction, which takes place at the rate τ, and $o(\Delta t)$ denotes terms that are second order in Δt. The growth rate of the economy in the stationary equilibrium can then be computed as

$$g = \eta\tau + \eta\xi\left[\sum\nolimits_{n=0}^{n^*-1} \mu_n^H \alpha^{n+1} + \sum\nolimits_{n=0}^{\infty} \mu_n^L \alpha^{n+1}\right]. \tag{17}$$

2.7 Equilibrium With $\kappa < 1$

In this subsection, we turn to the general case with $\kappa < 1$. We will show that the structure of the equilibrium is similar, except that now the switch to radical innovation for high-type firms will depend both on their current productivity and on their prior incremental innovations.

The value of a product line operated by low- and high-type firms can now be written, respectively, as:

$$\begin{aligned} rV_L(q_j,n) - \dot{V}_L(q_j,n) &= \max_a \{\pi q_j + \bar{q}_t f(a) - w_{a,t}\} + \xi\left[V_L\left(q_j + \eta_{n+1}, n+1\right) - V_L(q_j,n)\right] \\ &\quad -\tau V_L(q_j,n) + \varphi\left[V_H(q_j,n) - V_L(q_j,n)\right], \\ &\quad \text{and} \\ rV_H(q_j,n) - \dot{V}_H(q_j,n) &= \max\left\{ \begin{array}{c} \pi q_j + \max_a \left\{\bar{q}_t f(a) - w_{a,t} + \xi \left[\begin{array}{c} V_H\left(q_j + \eta_{n+1}, n+1\right) \\ -V_H(q_j,n) \end{array} \right]\right\}; \\ \pi q_j + \max_{a\geq 0}\{\bar{q}_t f(a) + \Lambda\theta_H \bar{q}^a \mathbb{E}V_H(t) - w_{a,t}\} \end{array} \right\} \\ &\quad -\tau V_H(q_j,n) + \psi\Lambda\theta_H \mathbb{E}V_H(t). \end{aligned}$$

Here note that, with a slight abuse of notation, we wrote $\mathbb{E}V_H(t)$ instead of $\mathbb{E}V_H(\bar{q}_t)$ for the value of a new radical innovation, since this depends in general not just on average current productivity in the economy, $\bar{q}_t$, but also on the distribution of product lines across different states. All the same, in the stationary equilibrium it will clearly grow at the same rate as $\bar{q}_t$, g. Second, η_n is now a function of both the current productivity of the firm and the average current productivity in the economy, $\bar{q}_t$.

With an argument similar to that in the previous subsection, the equilibrium wage schedule for managers will be given by

$$w_{a,t} = \begin{cases} f(a)\,\bar{q}_t & \text{for} \quad a > a^* \\ f(a)\,\bar{q}_t + \Lambda\theta_H\left[\bar{q}^a - \bar{q}^{a^*}\right]\mathbb{E}V_H(t) & \text{for} \quad a \leq a^* \end{cases}$$

This enables us to write simplified versions of the value functions as:

$$\begin{aligned} rV_L(q_j,n) - \dot{V}_L(q_j,n) &= \pi q_j + \xi\left[V_L\left(q_j + \eta_{n+1}, n+1\right) - V_L(q_j,n)\right] \\ &\quad -\tau V_L(q_j,n) + \varphi\left[V_H(q_j,n) - V_L(q_j,n)\right] \end{aligned}$$

and

$$\begin{aligned} rV_H(q_j,n) - \dot{V}_H(q_j,n) &= \max\left\{ \begin{array}{c} \pi q_j + \xi\left[V_H\left(q_j + \eta_{n+1}, n+1\right) - V_H(q_j,n)\right]; \\ \pi q_j + \Lambda\theta_H\bar{q}^{a^*}\mathbb{E}V_H(t) \end{array} \right\} \\ &\quad -\tau V_H(q_j,n) + \psi\Lambda\theta_H\mathbb{E}V_H(t). \end{aligned}$$

Proposition 5 *Consider the economy with $\kappa < 1$. Then, for a product line with current quality q operated by a high-type firm, the manager will be younger and will pursue radical innovation when the number of prior incremental innovations is greater than or equal to $n_t^*(q)$, where $n_t^*(q)$ is increasing in q. That is, a high-type firm is more likely to pursue radical innovation when its current productivity is lower and the number of its prior innovations in the same cluster is higher.*

Proof. See the Appendix. ■

This proposition thus establishes that in this generalized setup (with $\kappa < 1$), the main results from Proposition 4 continue to hold, but in addition we obtain the result that radical innovation is more likely when a high-type firm has lower current productivity (conditional on its prior number of incremental innovations); or conversely, for a given level of productivity, it is more likely when there has been a greater number of prior incremental innovations. We will investigate this additional implication in our firm-level analysis.[17]

[17] This result is related to the idea of "disruptive innovations" proposed in Christensen's *The Innovator's Dilemma* (1997). This result also clarifies that our potential answer to the innovator's dilemma, consistent both with Arrow's replacement effect and the results presented below, is that successful firms with higher sales have more to fear from disruptive innovations and tend to retrench and become less open to new ideas, practices and innovations.

3 Data and Variable Construction

In this section, we describe the various datasets we use and our data construction. We also provide some basic descriptive statistics.

3.1 Data Sources

USPTO Utility Patents Grant Data (PDP) The patent grant data are obtained from the NBER Patent Database Project (PDP) and contain data for all 3,279,509 utility patents granted between the years 1976-2006 by the United States Patent and Trademark Office (USPTO). This dataset contains extensive information on each granted patent, including the unique patent number, a unique identifier for the assignee, the nationality of the assignee, the technology class, and backward and forward citations in the sample up to 2006. Following a dynamic assignment procedure, we link this dataset to the Compustat dataset, which we next describe.[18]

Compustat North American Fundamentals We draw our main sample from the Compustat data for publicly traded firms in North America. This data set contains balance sheet items reported by the companies annually between 1974 and 2006. It contains 29,378 different companies, and 390,467 *company* $\times$ *year* observations. The main variables of interest are net sales, employment, firm age (defined as time since entry into the Compustat sample), SIC code, R&D expenditures, total liabilities, net income, and plant property and equipment as a proxy for physical capital.

Executive Compensation Data (Execucomp) Standard and Poor's Execucomp provides information on the age of the top executives of a company starting from 1992. We use information on CEO age or the average age of (top) managers of a company to construct proxies for openness to disruption at the firm level.[19]

The Careers and Co-Authorship Networks of U.S. Patent Inventors Extensive information on the inventors of patents granted in the United States between years 1975-2008 is obtained from Lai et. al.'s (2009) dataset. These authors use inventor names and addresses as well as patent characteristics to generate unique inventor identifiers upon which we heavily draw. Their dataset contains 8,031,908 observations at the *patent* $\times$ *inventor* level, and 2,229,219 unique inventors, and can be linked to the PDP dataset using the unique patent number assigned by the USPTO.

Cross-Country Data on Manager Age We also collected data on the age of the CEOs and CFOs of the 25 largest listed companies for 37 countries. We selected the top 25 companies,

[18] Details on the assignment procedure are provided at https://sites.google.com/site/patentdataproject/.

[19] We drop observations where reported CEO age is less than 26.

when available, according to the Financial Times' FT-500 list, which ranks firms according to their market capitalization. We completed the list by using information from transnationale.org when the FT-500 did not include 25 companies for a country. We then obtained the age of the CEOs and CFOs from the websites of the companies. Overall, our dataset has on average 20 companies and 31.6 managers (CEO or CFO) per country.

National Culture Dimensions The Dutch social scientist Geert Hofstede devised five different indices of national culture: power distance, individualism vs. collectivism, masculinity vs. femininity, uncertainty avoidance, and long-term orientation. The initial survey was conducted among IBM employees in 30 countries to understand cross-country differences in corporate culture. This work has been expanded with additional surveys that have been answered by members of other professions and expanded to 80 countries (see Hofstede, 2001, and Hofstede et al., 2010).[20] We use the individualism and uncertainty avoidance measures below.

The individualism measure is defined as "a preference for a loosely-knit social framework in which individuals are expected to take care of themselves and their immediate families only." A low individualism score is indicative of a more collectivist society, where social safety networks are more common and individuals are influenced by collective goals and constraints.

The uncertainty avoidance measure expresses the degree to which the members of a society seek to avoid uncertainty and ambiguity. Countries with a higher score are more rigid in terms of belief and behavior and are more intolerant of unorthodox ideas. On the other end of the spectrum, societies with a low score are more welcoming to new ideas and value practice above principles. Both the individualism and the uncertainty avoidance indices are normalized to lie between 0 and 1.

Other Data Sources We use the average years of schooling in secondary education from the Barro-Lee dataset as a proxy of the human capital of a country.[21] We also use real GDP per capita numbers and R&D intensity from the World Bank's World Development Indicators database.

In our baseline analysis, we focus on a *balanced panel* of firms, with complete information on all variables used in our cross-sectional analysis. To maximize the number of observations in this balanced panel, we focus on the years between 1995 and 2000, so citation and patents information in our baseline results come from 1995-2000 (with patents classified according to their year of *application*). We then extend our analysis to an *unbalanced panel* spanning 1992-2004 (we cannot go earlier than 1992 because our manager age data start at the state, and we cannot go further than

[20] http://geert-hofstede.com/national-culture.html

[21] http://www.barrolee.com/data/dataexp.htm. See Barro and Lee (2013) for details.

2004 as we need a subsequent window during which to measure citations). We also use citations from 1995-2000 in our cross-country analysis.

3.2 Variable Construction

Innovation Quality Our baseline measure of innovation quality is the number of citations a patent received as of 2006. We also use the truncation correction weights devised by Hall, Jaffe, and Trajtenberg (2001) to correct for systematic citation differences across different technology classes and also for the fact that earlier patents will have more years during which they can receive citations (we also experiment with counting citations during a five-year window for each patent). Based on this variable, an average innovation quality variable is generated at the *company* × *year* and *country* × *year* levels. For our cross-country dataset, the country of the assignee is used to determine the country to which the patent belongs.

Superstar Fraction A superstar inventor is defined as an inventor who surpasses his or her peers in the quality of patents generated as observed in the sample. A score for each unique inventor is generated by calculating the average quality of all the patents in which the inventor took part. All inventors are ranked according to this score, and the top 5% are considered to be superstar inventors. The superstar fraction of a company or country in a given year is calculated as the fraction of patents with superstar inventors in that year (if a patent has more than one inventor, it gets a fractional superstar designation equal to the ratio of superstar inventors to the total number of inventors of the patent). The country of the inventor is determined according to the country of the patent assignee.

Tail Innovations The tail innovation index is defined as the fraction of patents of a firm or country that receive more than a certain number of citations (once again using the truncation correction weights of Hall, Jaffe and Trajtenberg, 2001). Namely, let $s_{ft}(p)$ denote the number of the patents of a firm (or country) that are above the p^{th} percentile of the year t distribution according to citations. Then, the tail innovation index is defined as

$$\text{Tail}_{ft}(p) = \frac{s_{ft}(p)}{s_{ft}(0.50)},$$

where $p > 0.50$. This is of course also equivalent to the ratio of the number of patents by firm f at time t with citations above the p^{th} percentile divided by the number of patents by firm f at time t with citations above the median (and is not defined for firms that have no patents with citations above the median). For our baseline measure of tail innovations, we choose $p = 0.99$, so that our measure is the fraction of patents of a firm or country that are at the 99th percentile of citations divided by the fraction of patents that are at the median of citations. The reason we include $s_{ft}(0.50)$ in the denominator is that we would like to capture whether, controlling for "average" innovation output, some companies, innovators or countries have the tendency for generating "tail innovations" with very high citations.[22]

Generality and Originality We also use the generality and originality indices devised by Hall, Jaffe and Trajtenberg (2001). Let $i \in I$ denote a technology class and $s_{ij} \in [0, 1]$ denote the share of citations that patent j receives from patents in technology class i (of course with $\sum_{i \in I} s_{ij} = 1$). Then for a patent j with positive citations, we define

$$\text{generality}_j = 1 - \sum_{i \in I} s_{ij}^2.$$

This index thus measures the dispersion of the citations received by a patent in terms of the technology classes of citing patents. Greater dispersion of citations is interpreted as a sign of greater generality. The originality index is defined similarly except that we use the citations it gives to other patents. Both indices are averaged across all of the patents of a firm or a country to obtain our firm-level and cross-country generality and originality indices. The patent classes used are the 80 two-digit International Patent Classification (IPC) classes.

3.3 Descriptive Statistics

Panel A of Table 1 provides descriptive statistics for our balanced firm, unbalanced firm and cross-country samples. Since we focus on regressions weighted by the number of patents held by a company or country, all statistics are weighted by the number of patents. We multiply our indices for tail innovation, superstar fraction, generality, and R&D intensity by 100.

The table shows that average manager age is 52.3 in our firm-level (balanced or unbalanced Compustat) sample and 56.1 in our cross-country sample, while average CEO age is 55.3 in the balanced sample and 55.5 in the unbalanced sample. The comparison of our average number of citations per patent, superstar fraction, tail innovation, and generality indices shows that, as expected, our Compustat firms have higher values than the average country.

Panels B and C present the firm-level and cross-section country correlations between our main measures of creativity of innovations, which are quite highly correlated except for the generality index at the firm level. Panel D of Table 1 presents the correlation between our three cross-country indices of openness to disruption. These three indices are also highly correlated.

[22] If we do not include the correction in the denominator, all of the results reported below continue to hold, and are in fact stronger. When we turn to patent-level regressions, we will not be able to include such a correction (since there are no other patents that can be included in the denominator).

4 Firm-Level and Patent-Level Results

Our main empirical results exploit firm-level variation in manager age across Compustat companies. Recall that in our theory manager age is in part an indicator of a corporate culture that is open to disruption (because high-type firms that have a competitive advantage in radical innovation select to hire younger managers). But there is also a causal effect of manager age on creative innovations since, conditional on being employed by such a firm, a young manager contributes to radical/creative innovations (because of her more recent knowledge stock). Motivated by this reasoning, in this section we start with the cross-sectional relationship between firm-level measures of creative innovations and manager age (emphasizing throughout that our estimates do not necessarily correspond to the causal effect of manager age on creative innovations).[23] We then turn to a more direct investigation of the effect of manager age on creative innovations, focusing on regressions that exploit "within-firm" variation, and also investigate the timing of increases in creative innovations at the firm level and the relationship of this to our structural parameters.

4.1 Baseline Results

Our baseline results are presented in Table 2. Our estimating equation is

$$y_f = \alpha m_f + \mathbf{X}'_f \boldsymbol{\beta} + \delta_{i(f)} + \varepsilon_f, \tag{18}$$

where y_f is one of our measures of creative innovations introduced in the previous section (innovation quality, superstar fraction, tail innovation, or generality) for firm f, and m_f is our firm-level measure of openness to disruption, the average age of company CEOs over our sample window. In addition, $\mathbf{X}_f$ is a vector of controls, in this case, firm age, log of employment, log of sales, and log of total number of patents during our time window (we do not have measures of the human capital of the firm's employees).[24] Controlling for firm age is particularly important to distinguish the correlation of creativity of innovations with manager age from its correlation with firm age. In addition, $\delta_{i(f)}$ denotes a full set of four-digit main SIC dummies, so that the comparisons are always across firms within a fairly narrow industry.[25] Finally, ε_f is the error term.

Our baseline sample comprises 279 firms with complete information on CEO and positive patents between 1995 and 2000 (as well as information on firm age, sales, and employment). As noted above, we first exploit only cross-sectional information, so our regressions have one observation per firm,

[23] Another caveat is that our theoretical results relate manager age at the product-line level to the innovation strategy and creativity of innovations, while the bulk of our empirical analysis in this section will be at the firm level focusing on the age of a firm's CEO (or top managers).

[24] Our log employment and log sales variables in the cross-sectional regressions are computed as averages of annual log employment and log sales.

[25] All firms in our baseline sample are in one of 120 four-digit SIC industries.

and are weighted with the total patent count of the firm, so that they put less weight on observations for which our measures of creative innovations are computed from only a few patents. All standard errors are robust against heteroscedasticity. Different columns of Table 2 correspond to our four different measures of creative innovations.

Column 1 shows an economically sizable correlation between CEO age and our measure of innovation quality (average number of citations per patent). The coefficient estimate, -0.278 (standard error = 0.088), is statistically significant at 1% and indicates that companies with a younger CEO have greater innovation quality. We interpret this pattern as evidence that companies that are more open to disruption tend to be the ones producing more creative innovations. The quantitative magnitudes are significant and also plausible. For example, a one-year increase in CEO age is associated with a 0.278 increase in average citations, which is approximately 1.3% of the firm-level weighted mean of our innovation quality variable (20.5).

The pattern of the covariates is also interesting. Firm age is negatively associated with innovation quality, suggesting that younger firms are more creative (though this pattern is not as robust as the impact of CEO age in other specifications). Our measures of creative innovations are also uncorrelated with employment and sales, and are largely uncorrelated with the number of patents held by the firm (the exception being a marginally significant relationship for tail innovations). This confirms that our measures of creativity of innovations are quite distinct from the total number of patents.

Column 2 shows a similar relationship with the superstar fraction (-0.300, standard error = 0.141). This also suggests that younger CEOs tend to work with higher-quality innovators (a relationship we directly investigate in Table 9 below). Columns 3 and 4 show even more precisely estimated (significant at 1% or less) and economically large relationships with our measures of tail innovations and generality. The implied quantitative magnitudes are a little larger with the superstar fraction and tail innovation measures (a one-year increase in CEO age is associated with, respectively, 2.4% and 5.5% increases relative to weighted sample means in these two measures).

Overall, these results suggest that there is a strong statistical and quantitative relationship between the age of the CEO of a Compustat company and each one of our four measures of creative innovations. We next investigate the robustness of these patterns.

4.2 Robustness

Tables 3 and 4 probe the robustness of our firm-level results in different dimensions. Table 3 looks at the alternative measures of creative innovations (these are a measure of innovation quality using average citations per patent computed using only five years of citations data, a measure of superstar inventors using information on the most highly cited patent of the inventor, the tail innovation index with $p = 0.90$, and the originality index). The results show that the pattern is quite similar to those in Table 2, except that the relationship is no longer statistically significant with the alternative measure of the superstar fraction.

Table 4 looks at several different robustness exercises. Panel A replaces the four-digit SIC dummies with three-digit dummies (a total of 84 in our baseline sample), with effects very similar to our baseline results.

Panel B goes in the opposite direction and enriches the set of controls. In particular, this specification, in addition to the four-digit SIC dummies, includes several other firm-level controls: profitability (income to sales ratio), debt to sales ratio, and log physical capital of the firm. The results are virtually the same as those in Table 2, but a little more precisely estimated. For example, CEO age is statistically significant at less than 1% with all of our measures of creative innovations, except for the superstar fraction, for which it is significant at 5%.

Panel C, in addition, adds R&D intensity (R&D to sales ratio) to the previous specification.[26] This is intended to verify that our results cannot be explained by some firms performing more R&D than others (here the sample declines to 257 companies). The results are once again very similar to those in our baseline regressions in Table 2.

Panel D uses the average age of the top management team rather than CEO age. We prefer CEO age as our baseline measure because across companies there is considerable variation in the number of managers for which age data are available, making this measure potentially less comparable across firms. Nevertheless, the relationship is very similar to this measure as shown in Panel D.

Panels E and F reestimate the specifications in Table 2 for subsamples of high-tech and low-tech firms, where high-tech firms are those in SIC 35 and 36 (industrial and commercial machinery and equipment and computer equipment; and electronic and other electrical equipment and components), and low-tech firms are the rest. This is intended to check whether our results are driven by a subset of firms and whether they are differential between these two subsamples. The results are fairly similar in these two subsamples, except for the superstar fraction variable, which shows a considerably stronger relationship for the low-tech sample.

4.3 Panel Results

We now show that, though naturally much noisier, the correlation between CEO age and creative innovations is present when we exploit within-firm variation in the age of the CEO. We will also document, however, that consistent with our theory, creative innovations start increasing before there is a decline in CEO age.

With this objective in mind, in Table 5 we start with our baseline balanced sample, but now we

[26] To deal with outliers in R&D expenditures, we winsorize this variable at its 99th percentile value.

compute our measures of creativity of innovations at an annual frequency. The covariates we use are also at an annual frequency and include a full set of year dummies. As a first step, in Panel A, we maintain our key right-hand-side variable, average CEO age over the sample period, which is thus held constant across years in this panel. In this table, standard errors are robust for arbitrary heteroscedasticity at the firm level (thus allowing for arbitrary dependence across the observations for the same firm). These specifications are directly comparable to those in Table 2, and indeed, the coefficient estimates and standard errors are very similar (though they are not identical since the covariates are now time-varying).[27]

Panel B extends our sample in two different ways. First, we include firms that were left out of the balanced panel (i.e., firms for which CEO age or patent information is available in some but not all years). Second, with the unbalanced panel, we can now consider a longer sample spanning $1992-2004$ (we cannot go before 1992 because of lack of data on manager age, and we prefer not to go beyond 2004, as this would make the citation window too short and thus our measures much less reliable). The resulting sample has 7111 observations (or 5803 observations with tail innovation, since we lose firm-years when no patent is above the median of the citation distribution). Despite the increase in the number of firms to 1256 (from 279) and the addition of seven more years of data, the results are remarkably similar to those in Panel A and to our baseline estimates.

Panel C allows CEO age to vary across years but also includes firm fixed effects as well as year effects (and, of course, in this case, SIC industry dummies and firm age are dropped). This effectively means that the CEO age variable is being identified from changes in CEOs.[28] Hence, this is a demanding specification investigating whether in years when a firm has a younger CEO, it tends to have more creative innovations, and this motivates our choice of focusing on the $1992-2004$ sample for this exercise. In addition to throwing away all of the between-firm variation, another challenge to finding meaningful results in this specification is that patent applications in one year are often the result of research and product selection from several past years.[29] Though these considerations stack the cards against finding a significant relationship between CEO age and creative innovations, the results are generally quite consistent with our cross-sectional estimates from the balanced panel. All of the coefficient estimates in these within-firm regressions, except for generality, have the same sign and are statistically significant as in our baseline results in Table 2. For innovation quality, the magnitude of the estimate is about 12% larger than the specification

[27] The number of observations is now lower in columns 3 and 4 because not all firms have patents with citations above the median (for tail innovations) or with positive citations (for generality) in all years.

[28] This specification is related to Bertrand and Schoar's famous (2003) paper on the effect of managers on corporate policies though, in contrast to our focus on CEOs, their sample includes chief financial and operating officers as well as lower-level executives.

[29] Recall, however, that patents are classified according to their year of *application*, so we are investigating the impact of CEO age not on patents granted when the CEO is in charge but on patents applied for when the CEO is in charge.

without fixed effects in Panel B (e.g., -0.188 vs. -0.168), whereas for superstar fraction and tail innovations, it is smaller—about 47% to 73% of the magnitude in Panel B.

The current CEO/manager influences the contemporaneous innovation strategy, and in our model, this has an immediate impact on radical innovations. In practice, some of the impact is likely to be delayed, since research projects, and even patenting, can take several years. We may therefore expect the impact of the CEO's human capital, decisions and age to influence the creativity of innovations over time. We investigate this issue in Panel D by including current CEO age and lagged CEO age simultaneously. Our results show that, with all of our measures of creative innovations (except generality), both matter with quantitatively similar magnitudes.

A related question concerns separating the impact of the current CEO from the persistent effects of past innovations—in particular, past creativity may spill over into current creativity in part because patents from the same project may arrive in the course of several years. We investigate this issue by including the lagged dependent variable on the right-hand side. Though such a model, with fixed effects and lagged dependent variable, is not consistently estimated by the standard within estimator when the coefficient on the lagged dependent variable is close to 1, the results in Panel E show that its coefficient is very far from 1 and the estimates are fairly similar to those in Panel C.[30]

Finally, in Panel F we turn to the more detailed longitudinal implications of our model—that creativity of innovations should increase, on average, before the firm switches to a younger manager. The simplest way of investigating this prediction is by including the lead of CEO age together with current CEO age (similar to the specification in Panel D, except that lead CEO age replaces lagged CEO age). The specifications reported in Panel F show statistically significant negative effects of both current and lead CEO age on the creativity of innovations (except with the generality measure). Interestingly, and perhaps somewhat surprisingly, the magnitudes of the lead and the contemporaneous effects are quite similar. The significant effect of lead CEO age is prima facie evidence of the importance of sorting of younger CEOs to firms that are firms that are more open to disruption (more likely to have creative innovations).

Although the results in Panel F suggest that both the sorting and the causal effects of CEO/manager age are important for the creativity of innovations, they do not directly translate into an estimate of the impact of CEO/manager age on creative innovations for a given firm (because changes in CEO age are associated with changes in firm type/corporate culture as well as the firm's prior number of incremental innovations). We next turn to an indirect inference procedure utilizing the structure of our model to obtain an estimate of the size of this causal effect.

[30] If we estimate these models using Arellano and Bond's (1991) GMM estimators, the results are similar with innovation quality and superstar fraction, but weaker with the tail innovation index, partly because we lose about a quarter of our sample with these GMM models.

4.4 The Causal Effect of Manager Age on Creative Innovations

In this subsection, we perform a simple indirect inference exercise in order to shed further light on the causal effect of manager age on creative innovations. We choose the parameters of the model presented in Section 2 so that the model quantitatively matches the reduced-form estimates—in particular, the coefficients of lead and current CEO age for innovation quality. We then use these implied parameters to compute the implied causal effect of manager age on creative innovations given these parameters.

The (average) impact of a younger manager on the creativity of innovations for a given firm type is $\frac{1}{F(a^*)}\int_0^{a^*}\Lambda\theta_H\bar{q}^a dF(a) = \frac{\Lambda\theta_H\delta}{g+\delta}\frac{\left[1-e^{-(g+\delta)a^*}\right]}{\left[1-e^{-\delta a^*}\right]}$. Because of the sorting of younger managers to firms that are more open to disruption, we cannot read off this quantity from our reduced-form empirical exercise. Rather, we need to obtain estimates of the parameters ψ and $\Lambda\theta_H$ (the parameters Λ and θ_H do not matter separately, and thus in what follows, we will treat $\Lambda\theta_H$ as a single parameter). The reduced-form coefficient estimates are functions of these parameters, but they also depend on the transitions between high-type and low-type firms, the distribution of incremental innovations per product relative to the threshold for radical innovation, n^*, and the stationary distributions theoretically characterized above.

Though structurally estimating all of the underlying parameters of our model would require more information on firm transitions and stationary distributions, we can obtain estimates of the structural parameters that are relevant for the extent of the causal effect of CEO age on creative innovations from a simple indirect inference exercise. For this exercise, we set the discount rate to $\rho = 0.02$, and normalize the profit flow to $\pi = 1$ (which is without loss of any generality). We fit an exponential distribution to the age distribution of managers in our sample to obtain an estimate of δ in the model. We take the entry rate to be $x = 5\%$ which corresponds to the entry rate in our Compustat sample. Finally, we take the parameter α, which determines how rapidly the productivity of incremental innovations declines from Akcigit and Kerr (2015), who estimate a similar parameter from the patent citation distribution.

This leaves the following parameter vector $\Psi \equiv \{\psi, \varphi, \Lambda\theta_H, \xi, \eta, \zeta\}$ to be determined. Once these parameter values have also been fixed, optimal innovation decisions and equilibrium stationary distributions can be computed using the expressions provided in Section 2. We can then generate simulated firm histories from which the equivalents of the reduced-form regression coefficients in Table 5 can be computed. Of particular importance for this exercise are the specifications in Panels C and F of Table 5, where various measures of creative innovations were regressed on current CEO age (and lead CEO age in Panel F), firm fixed effects and controls.

Let us denote the coefficient estimate on current CEO age in column 1, Panel C of Table 5 by $\gamma_{current}$, and the coefficient estimates on current and lead CEO age in column 1, Panel F, respectively, by $\gamma'_{current}$ and γ'_{lead}. In our indirect inference procedure, we will target these three parameters. Specifically, we generate data from the model given a parameter vector Ψ, and convert the measure of successful radical innovation in the model, which is a 0-1 variable, into the same units as our innovation quality variable (by dividing it by its variance and multiplying it with the variance of innovation quality). We then run the same regression as in Panel C and F of Table 5, and compare the estimates to the empirical estimates of $\gamma_{current}$, $\gamma'_{current}$ and γ'_{lead}.

In addition to these three regression coefficients, our indirect inference procedure targets three central moments in the data: the average annual growth rate of (real) sales per worker, within-firm coefficient of variation of radical innovations, and the fraction of incremental innovations, measured as fraction of internal patents which mainly build on innovating firm's existing lines (as opposed to innovating on product lines operated by other firms).[31] This implies that we have in total six data moments and six parameters.

Finally, we make two additional assumptions in matching the model to data. First, in the model managers are employed at the product line level, whereas in the data we only observe managers/CEO at the level of the company (which comprises several product lines). We ignore this distinction, and treat the data as if it were generated from one product firms. Second, in the model, the identity of the manager/CEO is indeterminate as there are no costs of changing managers, so a firm could change its manager every instant or at some regular interval even without changing its innovation strategy. To make the model more comparable to the data, we assume that a firm keeps its manager until it needs to switch from an older to a younger manager in order to change its innovation strategy.

Table 6 provides the values of the parameters we have selected on the basis of external data as well as the values of the parameters in the vector Ψ, which are chosen to match the six aforementioned moments. Table 7 shows the match between the values of these moments in the data and those implied by the model. The model-implied numbers are on the whole very close to the targeted empirical moments. The most important lesson from Table 7 is that the model is quite consistent with reduced-form regression results, including the significant and sizable coefficient on lead CEO age, which is generated by the fact that $\psi > 0$ and is a non-trivial source of creative innovations.

The implied pattern is also visible in Figures 1 and 2, which plot the probability of a creative innovation and the average CEO age as a function of time since switching to high-type. These figures show that firms slowly reduce the average age of their managers after switching to high-type (if at first they are below n^*, they do not need to change their CEO). Correspondingly, they

[31] Following Akcigit and Kerr (2015), we define internal patents as those whose majority of citations are self cites.

also slowly increase their probability of creative innovations. Because much of this increase in the probability of creative innovations takes place before firms switch to a younger manager, in the reduced-form regressions it will be captured by lead CEO age.

It is also useful to gauge whether, at these estimated parameter values, the model performs well on some other dimensions. One empirical moment we have not used for estimation is the probability of firms switching to younger managers. At the estimated parameter values, 6% of all firms attempt a radical innovation (these are high-type firms with $n \geq n^*$). Consequently, "young" managers (defined as those with $a < a^*$ in Proposition 4) also make up 6% of the population of managers, implying that a^* corresponds to age 43 in our sample of managers/CEOs. Using this information, we can then compare the annual probability of a firm switching from an old manager (with $a > a^*$) to a young manager (with $a \leq a^*$) in the data and in the model. Reassuringly, these two numbers are very close to each other, respectively 0.62% and 0.63%.

Using these parameter estimates, we next compute the "causal effect" of manager age on creative innovations. There are several ways in which such a causal effect might be defined. First, we could define the causal effect in a fashion analogous to "treatment effect on the treated," by considering the loss of creative innovation that firms that were previously hiring young managers and pursuing radical innovations would suffer. Second, one could focus on the "average treatment effect," corresponding to the impact of having a younger manager for an average firm in our sample. It is intuitive that these two measures of causal effects will be quite different, since, as just noted, only 6% of firms are attempting radical innovations in our stationary distribution.

For the first, we start with the equilibrium stationary distribution and reshuffle managers only among firms attempting radical innovation (which are high-types with more than n^* incremental innovations and hiring managers younger than a^*), and we repeat this for 13 periods. Because such firms will continue to attempt radical innovation after the reshuffling (since the reshuffling involves only managers younger than a^*), the change in the likelihood of radical innovation of any given firm captures the causal effect of a younger (or older) manager on a firm attempting radical innovation. We quantify this effect by running the same regression of the likelihood of radical innovation on the age of manager for this subsample of firms (corresponding to Panel C) and then comparing it to the reduced-form relationship between these two variables in the model-generated data, -0.211 (as reported in Table 7).[32] The resulting causal effect is estimated as -0.040. This effect is thus considerably smaller than the reduced-form regression coefficient of -0.211, though it should not be directly compared to this number, which applies to the entire sample, while the causal effect of -0.040 is only for 6% of the entire sample. To obtain a causal effect estimate more comparable to

[32] We use the regression coefficient obtained from model-generated data rather than the regression coefficient from Table 5, -0.188, since this will be compared to numbers also obtained from model-generated data.

the reduced-form relationship, we next turn to the average treatment effect.

For this second exercise, we again start with the stationary distribution and reshuffle managers randomly for 13 periods. We assume that after the reshuffling, each firm will pursue the same innovation strategy.[33] We then use the data generated by this thought experiment to run a regression of the likelihood of a radical innovation on manager age for the entire sample of firms (including both low-type firms and high-type firms not attempting a radical innovation). This exercise yields an average causal effect of -0.003, and thus accounts approximately for 1.5% ($\simeq 0.003/0.211$) of the relationship between CEO age and creative innovations. The rest of this relationship is explained by sorting effects—because it is high-type firms that are hiring younger managers.

It is intuitive that the first estimate of the causal effect is much larger than the second, because it explicitly focuses on the small subsample of firms attempting a radical innovation. But even in this case, especially once we take into account that this causal effect applies only to 6% of the sample of firms that are attempting a radical innovation, much of the association between manager age and creative innovations is accounted for by the sorting of younger managers to firms that are more open to disruption.

Overall, our indirect inference exercise establishes that the model can generate the patterns we see in the data, and implies that much of the reduced-form relationship between manager age and creative innovations is due to sorting, but also that there is a small causal effect of younger managers on creative innovations as well.

4.5 Inventor Age and Creativity of Innovations

We next turn to patent-level regressions to investigate the relationship between the age of inventor—defined as any inventor listed in our patent data—and our various measures of creativity of innovations. Though in our theoretical model there is no distinction between managers and inventors, this distinction is of course important in practice. One might then expect the role of product-line managers in our model to be played partly by the top management of the firm and partly by inventors (or the lead inventor) working on a particular R&D project. CEOs, then, not only decide which projects the company should focus on but also choose the research team. In this subsection, we bring in information on the age of inventors in order to investigate the effect of manager/inventor age on the creativity of innovations once we control for the type of characteristics of the firm.

[33] It is possible that some firms would switch their innovation strategy because they end up with much older or much younger managers. However, whether this is the case or not would also depend on managerial wages after reshuffling, which in turn depends on a variety of auxiliary assumptions on wage determination under "mismatch". Our strategy avoids this complication, but estimates a lower bound on this effect, though this lower bound is likely to be fairly tight since low-type firms cannot change their innovation strategy and most high-type firms would be unlikely to alter their innovation strategy unless there is a very large change in the age of the manager assigned to them.

We use Lai et. al.'s (2009) unique inventor identifiers described above to create a proxy for this variable. Our proxy is the number of years since the first innovation of the inventor, which we will refer to as "inventor age."

Our main regression in this subsection will be at the patent level and take the form

$$y_{ift} = \phi I_{ift} + \alpha m_{ft} + \mathbf{X}'_{ift}\boldsymbol{\beta} + \delta_f + \gamma_i + d_t + \varepsilon_{ift}. \tag{19}$$

Here y_{ift} is one of our measures of the creativity of innovation for (patent) i granted to firm f at time t. Our key right-hand-side variable is I_{ift}, the age of the inventors named in patent i (in practice, there is often more than one such inventor listed for a patent). In addition, m_{ft} is defined as CEO age at time t and will be included in some regressions, $\mathbf{X}_{ift}$ is a vector of possible controls, and δ_f denotes a full set of firm fixed effects, so that our specifications here exploit differences in the creativity of innovations of a single firm as a function of the characteristics of the innovators involved in the relevant patent. In our core specifications, we also control for a set of dummies, denoted by γ_i, related to inventor characteristics as we describe below. All specifications also control for a full set of year effects, denoted by d_t, and ε_{ift} is the error term.[34]

The results from the estimation of (19) are reported in Table 8. In Panel A we focus on a specification similar to the regressions with firm fixed effects reported in Table 5. This is useful for showing that this different frame still replicates the results showing the impact of CEO age on creativity of innovations. In particular, Panel A focuses on Compustat firms for the period $1992-2004$ and includes the same set of controls as in Table 5 Panel C (firm fixed effects, year fixed effects, log employment, log sales and log patents of the firm); it does not contain any variables related to inventor characteristics. As in the rest of this table, these regressions are not weighted (since they are at the patent level) and the standard errors are robust and clustered at the firm level.

Our results using this specification are similar to those of Panel C of Table 5, though a little smaller. In column 1, for instance, we see an estimate of -0.119 (standard error $= 0.038$) compared to -0.188 in Table 5. We cannot define our measure of the superstar fraction and tail innovations in these patent-level regressions. We can, however, look at a patent-level measure related to tail innovations, a dummy for the patent in question being above the pth percentile of the citation distribution. We report results using this measure for two values, $p = 0.99$ and $p = 0.90$, in columns 2 and 3. Both of these measures are negatively correlated with CEO age, though only marginally significantly in these specifications.[35]

[34] A single patent can appear multiple times in our sample if it belongs to multiple firms, but this is very rare and applies to less than 0.2% of the patents in our sample.

[35] For completeness, we also show results with the generality index, even though the results in Table 5 already indicated that, with firm fixed effects included, there is no longer a significant relationship between CEO age and

Panel B goes in the other direction and reports the estimates of a model that controls for inventor characteristics and looks at the impact of inventor age, without controlling for CEO age, for the same sample as in Panel A (thus restricting it to firms with information on CEO age). As with all of the other models reported in this table, in Panel B we control for a full set of dummies for the maximum number of patents of any inventor associated with the patent in question has over our sample period;[36] a full set of dummies for the size of the inventor team (i.e., how many inventors are listed); and a full set of dummies for the three-digit IPC class.[37] The inclusion of this rich set of dummy variables enables us to compare inventors of similar productivity. It thus approximates a model that includes a full set of inventor dummies.[38] The results show that there is a strong relationship between inventor age and the creativity of innovations. For example, in column 1, the coefficient estimate on inventor age is -0.234 (standard error $= 0.026$), about twice as large as the CEO age estimate in Panel A.

When we do not control for CEO age, the sample can be extended beyond $1992-2004$. This is done in Panel C, which expands the sample in two different ways, first by including Compustat firms without CEO information, and second by broadening the time period covered to $1985-2004$. The results are very similar to those in Panel B, indicating that the focus on Compustat firms with CEO age information is not responsible for the broad patterns we are documenting.

Panel D extends the sample further to non-Compustat firms, which can also be included in our analysis since we are not using information on CEO age. This increases our sample sixfold (since most patents are held by non-Compustat firms). However, in this case, we can no longer include the employment and sales controls. Despite the addition of almost 1.5 million additional patents and the lack of our employment and sales controls, the results in this panel are again very similar to those in previous panels, and suggest that, at least in this instance, our results are not driven by our focus on the Compustat sample.

Panel E provides our main results in this subsection. It returns to the Compustat sample over the period 1992-2002 and adds back the CEO age variable; otherwise, the specification is identical to that in Panel B. The results show precisely estimated impacts of both CEO age and inventor age. For example, in column 1 with our innovation quality variable, the coefficient on CEO age is -0.111 (standard error $= 0.038$) and that on inventor age is -0.235 (standard error $= 0.027$);

the generality index, and this lack of relationship persists for all of the estimates we report in Table 8 (and for this reason, though we do show them for completeness, we will not discuss them in detail).

[36] In other words, we include a dummy variable for the assignee/inventor of this patent with the highest number of total patents having $k = 1, 2, ..., 89+$ patents (where 89+ corresponds to 89 or more patents for the inventor with the maximum number of patents).

[37] This corresponds to 374 separate technology classes and is roughly at the same level of disaggregation as the SIC dummies we used in the firm-level analysis in Tables 2-4.

[38] We cannot include a full set of inventor fixed effects directly because inventor age would not be identified in this case since we also have a full set of year dummies.

these are very close to the estimates in Panels A and B, respectively. The pattern is similar in the other columns (except again for generality).

These results provide further evidence that the relationship between manager/CEO age and the creativity of innovations in the data reflects an important dimension of sorting. In particular, firms appear to make several associated changes—in top management and innovation teams—around the same time they change their portfolio of innovation and their innovation strategy (and perhaps their "corporate culture"). Reflecting this sorting, the estimated magnitudes linking CEO age to our indices of creative innovations are smaller in Table 8 than those in our baseline firm-level regressions. Our next results, reported in Table 9, provide some direct evidence on this by looking at the relationship between inventor age and CEO age. In particular, we estimate a regression similar to equation (19) except that now the dependent variable is the average age of the inventors on the patents granted for that firm in year t and the key right-hand-side variable is the age of the CEO, and firm fixed effects are again controlled for. The first column of Table 9 reports a regression of the average age of inventors on firm and year fixed effects, log employment, log sales, log patents, and CEO age, while the second column also adds dummies for inventor team size and three-digit IPC class as in the specifications in Table 8. The results, which show a positive (though only marginally significant) relationship, suggest that younger CEOs tend to hire younger inventors, indirectly corroborating the sorting effect emphasized in our theoretical model.[39]

4.6 Stock of Knowledge, Opportunity Cost and Creativity of Innovations

Finally, Table 10 turns to an investigation of some additional implications of our approach already highlighted in our theoretical model (in particular, Proposition 5). We noted there that we may expect openness to disruption to be more important for companies that are technologically more advanced (as measured by the number of patents), but also that companies that have more to lose (because of the greater opportunity cost of disruption in terms of other profitable activities) may shy away from disruptive creative innovations. The firm-level data enable us to look at this issue by including the interaction between CEO age and log total patent count (as a proxy for how advanced the technology of the company is) and also the interaction between CEO age and log sales (as a proxy for company revenues that may be risked by disruptive innovations). According to the theoretical ideas suggested above, we expect the interaction with log total patent count to be negative, and that with sales to be positive (indicating that average manager age matters more for the creativity of innovations for companies with a significant number of patents and less for companies with high sales).

This is a demanding, as well as crude, test, since neither proxy is perfect, and moreover, log

[39] Interestingly, this result disappears when we do not control for firm fixed effects.

sales and log patent counts are positively correlated (the weighted correlation between the two variables is 0.7 in our sample), thus stacking the cards against finding an informative set of results.

Nevertheless, Table 10, which uses the same unbalanced sample with annual observations as in Table 5 Panel C, provides some evidence that our theoretical expectations are partially borne out. In all of our specifications, the interaction between CEO age and log total patent count is negative and the interaction with log sales is positive. Moreover, these interactions are statistically significant except for the log patent interaction for the innovation quality measure.[40] These results thus provide some support for the hypothesis that the stock of knowledge of the company and opportunity cost effects might be present and might in fact be quite important (at least quantitatively at this correlational level).

5 Cross-Country Correlations

In this section, we provide evidence that the firm-level relationship between manager age and creativity of innovations appears to aggregate up to the national level. In particular, we document that there is a cross-country relationship between manager age and creativity of innovations. Moreover, at the cross-country level, we can also use other indices potentially proxying for openness to disruption, which also show similar results, thus partially corroborating our interpretation of the manager age variables in our firm-level and cross-country empirical work.

The interpretation of the cross-country relationships should be somewhat different than the firm-level ones. At the country level, manager age, like our other measures of openness to disruption presented below, is likely to have its impact on the creativity of innovations not just because of its association with— and because of its impact on—firm-level innovation strategies, but also through economy-wide institutions, attitudes and values of the society. This suggests that the quantitative magnitudes of the relationships might be somewhat larger at the country level than at the firm level.

Our main cross-country results are presented in Table 11, which reports OLS regressions of the following form:

$$y_c = \alpha I_c + \mathbf{X}_c' \boldsymbol{\beta} + \varepsilon_c, \tag{20}$$

where y_c is one of our measures of creative innovations (innovation quality, superstar fraction, tail innovation, or generality) for country c, I_c denotes one of our measures of openness to disruption (the individualism index, the uncertainty avoidance index, or average manager age), $\mathbf{X}_c$ is a vector of controls (including average log real GDP per capita of the country, average years of secondary schooling and log of total patents of the country during this time period), and finally, ε_c is an

[40] As noted above, the main effects are evaluated at the sample mean and are typically close to the estimates reported in Table 2.

error term.[41] The coefficient of interest is α, which will reveal whether there is a cross-country correlation between our measures of openness to disruption and the creativity of innovations.

All regressions in Table 11 include one observation per country. As with our firm-level specifications, these regressions are weighted using the total number of patents as weights, which is again motivated by the fact that countries with more patents are both more important for their contribution to creative innovations and have much more precisely estimated measures for our key variables (see Appendix Table A1 for the distribution of total number of patents across countries). All standard errors continue to be robust against heteroscedasticity.

Panel A of Table 11 focuses on our measure of manager age (which is available for 37 countries). The patterns are very similar to those we obtained in the firm-level analysis, and show a strong correlation between average manager age and all four of our measures of creative innovations. For example, in column 1, the estimate of α is -0.484 (standard error = 0.225). We also see that log GDP per capita and average years of secondary schooling are not significant correlates of the creativity of innovations, while log patent count is significant and indicates that countries that have more patents also tend to have more creative innovations. Consistent with the caveat about the interpretation of the cross-country results, the quantitative magnitudes are somewhat larger than the firm-level ones: a one-year change in manager age increases average citations by 0.48 (3.3% compared to its mean of 14.5), the superstar fraction by 0.96 (14.4% relative to its mean), tail innovations by 0.23 (11.7% relative to its mean) and generality by 0.28 (1.3% relative to its mean). These effects are about 2 to 5 times larger than the firm-level estimates presented above.[42]

Panel B is similar to Panel A, except that it uses Hofstede's individualism index (this increases the sample from 37 to 50). The results are very similar to those using average manager age, and the quantitative magnitudes of the correlation between individualism and innovation quality are again sizable and somewhat larger than those in Panel A.[43]

Panel C has exactly the same structure, except that the right-hand-side variable is Hofstede's uncertainty avoidance index. The patterns are very similar and generally even more precisely esti-

[41] An additional covariate that might be useful to control for would be the average educational attainment of managers in a country. Though this number is available in the World Bank dataset that Gennaioli et al. (2013) use, there is very little overlap between this developing country sample and ours. We have instead experimented with controlling for the average education of the managers of the companies we have used for compiling our average manager age variable. This has no effect on the results reported here and is omitted to save space. The details are available upon request from the authors.

[42] If, instead, we look at the quantitative implications of moving from the 75th percentile of the manager age distribution to the 25th percentile, the magnitudes are more similar to the firm-level estimates. For example, moving from the country at the 25th percentile of average manager age in our sample to the 75th percentile (from 51.5 to 54.3) reduces our measure of innovation quality by 9.4% relative to the sample mean (14.5).

[43] For example, moving from the country at the 25th percentile of individualism in our sample to those at the 75th percentile (from 0.19 to 0.73) increases our measure of innovation quality by 19% relative to its weighted sample mean (14.5). Using the same metric for quantitative magnitude for the average manager age gives an increase in innovation quality by 9.4% relative to the sample mean (14.5).

mated (though, of course, they are now negative, since greater uncertainty avoidance corresponds to less openness to disruption). The quantitative magnitudes are similar to those in Panel B.[44]

Tables 12 and 13 probe the robustness of these cross-country relationships. Table 12 looks at various alternative measures of creative innovations (which we also investigate at the firm level). These are average citations per patent but now constructed using only a five-year window (so that we do not have to rely on the correction factors); an alternative measure of the superstar fraction of patents but now computed using information on the most highly cited patent to the inventor (rather than lifetime average citations); the tail innovation index computed with $p = 0.90$ (instead of $p = 0.99$); and the originality index mentioned above. The results in all cases are similar to the baseline (though weaker and not statistically significant with the alternative measure of superstar fraction).

Table 13 , on the other hand, investigates whether these results can be explained by the fact that R&D intensity (defined as total R&D spending divided by GDP at the country level) differs across countries. Our results largely might be reflecting the fact that some countries invest more in R&D and as a result generate more creative innovations. However, in our sample R&D intensity is not systematically related to individualism, uncertainty avoidance, or average manager age. Moreover, Table 13 shows that controlling for variation in R&D intensity does not change the basic correlations in our sample. The parameter estimates do change in some cases, particularly with the individualism variable, but the association between our measures of openness to disruption and creativity of innovations always remains highly significant.[45]

6 Conclusion

Despite a large theoretical and now a growing empirical literature on innovation, there is relatively little work on the determinants of the creativity of innovative activity, and in particular, the likelihood of innovations and patents that contribute most to knowledge. In this paper, building on Schumpeter's ideas, we suggested that openness to new ideas, disruptive innovations and unconventional practices—which we called openness to disruption, for short— may be a key determinant of creative innovations, and likewise, resistance to such disruptive behavior may hold back some of the most creative innovative activities.

[44] We do not run regressions including multiple indices at the same time, since we believe this type of horse race would not be particularly informative. Instead, we interpret each of these variables as a proxy for the same underlying tendency for openness to innovation, new practices and ideas.

[45] We also experimented with using cross-country differences in demographics to instrument for average manager age differences. Though these results corroborate the patterns shown here, we do not report them both because demographics could have a direct effect on the creativity of innovations, invalidating the exclusion restriction of such a strategy, and because we view the cross-country results as additional evidence rather than as our main empirical focus.

We provided a simple model drawing a clear distinction between radical (more creative) innovations and incremental innovations, whereby the former combines ideas from several different lines of research and creates more significant advances (and contributions to knowledge). These advances can be discouraged or even stopped, either through pecuniary or non-pecuniary means, preventing radical innovations directly or discouraging cross-fertilization of ideas from different fields.

The bulk of our paper provides illustrative cross-country and firm-level correlations consistent with the role of openness to disruption. We use several measures to proxy for creative innovations. These include our proxy for innovation quality, which is the average number of citations per patent; two indices for creativity of innovations, which are the fraction of superstar innovators and the likelihood of a very high number of citations (in particular, tail citations relative to median citations); and the generality index.

Our main proxy for openness to disruption is the age of the CEO or top management of the company (or the average age of the CEO and CFO of the top 25 publicly listed companies in a country). This variable is motivated based on the idea that only companies or societies open to such disruption will allow the young to rise up within the hierarchy. This is the only variable we have available as a proxy for openness to disruption at the firm level. At the country level, we augment this variable with the popular indices for individualism and uncertainty avoidance based on the work by the Dutch social scientist Geert Hofstede.

Using these proxies, at the firm, patent and country level, we find fairly consistent and robust correlations between openness to disruption and creative innovations. We also show that these relationships are generally robust. They do not, however, correspond to the causal effect of CEO (or manager) age on creative innovations because, as highlighted by our theoretical model, younger managers tend to be employed by firms that are more open to disruption and more creative. A simple indirect inference exercise using the structure of our model suggests that most of the empirical relationship between CEO age and creative innovations is due to these sorting effects, and the causal impact of CEO age is quite small.

Finally, our theoretical model further suggests that the impact of openness to disruption should be larger for companies that are technologically more advanced (closer to the technology frontier) and smaller for companies that have a greater opportunity cost of disruptive innovation. The empirical patterns in our firm-level data support this prediction.

We view our paper as a first step in the study of the impact of various social and economic incentives on creative activities and, in particular, on creative innovations. Future work investigating the causal effect of manager age on creative innovations using more systematic structural estimation techniques is an obvious next step. Further study of various other firm-level or cross-country characteristics on the creativity of innovations is also a natural direction. Another fruitful direction would be to systematically investigate what types of firms and firm organizations encourage creativity and lead to more creative innovations. This would involve both theoretical and empirical analyses of the internal organization of firms and their research strategies and a study of the interplay between institutional and society-level factors and the internal organization of firms.

Appendix: Omitted Proofs

Proof of Proposition 2. We conjecture that the value function for low-type firms takes the form in (10). Substituting this conjecture into (7), we get

$$\begin{aligned} r\left[Aq_j + B\left(n\right)\bar{q}_t\right] - B\left(n\right)g\bar{q}_t &= \pi q_j + \xi A\bar{q}_t\eta\alpha^{n+1} + \xi\left[B\left(n+1\right)\bar{q}_t - B\left(n\right)\bar{q}_t\right] \\ &\quad -\tau Aq_j - \tau B\left(n\right)\bar{q}_t + \varphi\left[Aq_j + \bar{q}_t\tilde{B}\left(n\right) - Aq_j - B\left(n\right)\bar{q}_t\right]. \end{aligned}$$

Equating the coefficients on q_j and $\bar{q}_t$, we obtain

$$rAq_j = \pi q_j - \tau Aq_j,$$

and

$$rB\left(n\right)\bar{q}_t - B\left(n\right)g\bar{q}_t = \xi A\bar{q}_t\eta\alpha^{n+1} + \bar{q}_t\xi\left[B\left(n+1\right) - B\left(n\right)\right] - \tau B\left(n\right)\bar{q}_t + \bar{q}_t\varphi\left[\tilde{B}\left(n\right) - B\left(n\right)\right].$$

Solving these equations for A and $B\left(n\right)$, while taking $\tilde{B}\left(n\right)$ as given and to be determined in Proposition 3, completes the proof. ■

Proof of Proposition 3. Following the same steps, we conjecture that the value function for high-type firms takes the form in (12), and substitute this into (8) to get

$$\begin{aligned} r\left[Aq_j + \bar{q}_t\tilde{B}\left(n\right)\right] - g\bar{q}_t\tilde{B}\left(n\right) &= \max\left\{\begin{array}{c} \pi q_j + \xi\left[A\bar{q}_t\eta\alpha^{n+1} + \bar{q}_t\tilde{B}\left(n+1\right) - \bar{q}_t\tilde{B}\left(n\right)\right]; \\ \pi q_j + \Lambda\theta_H\bar{q}^{a^*}\left[A\bar{q}_t + A\eta\bar{q}_t + \bar{q}_t\tilde{B}\left(0\right)\right] \end{array}\right\} \\ &\quad +\psi\Lambda\theta_H\left[A\bar{q}_t + A\eta\bar{q}_t + \bar{q}_t\tilde{B}\left(0\right)\right] - \tau\left[Aq_j + \bar{q}_t\tilde{B}\left(n\right)\right], \end{aligned}$$

which implies

$$\begin{aligned} \left(r+\tau\right)\left[Aq_j + \bar{q}_t\tilde{B}\left(n\right)\right] - g\bar{q}_t\tilde{B}\left(n\right) &= \pi q_j + \psi\Lambda\theta_H\left[A\bar{q}_t + A\eta\bar{q}_t + \bar{q}_t\tilde{B}\left(0\right)\right] \\ &\quad + \max\left\{\begin{array}{c} \bar{q}_t\xi\left[A\eta\alpha^{n+1} + \tilde{B}\left(n+1\right) - \tilde{B}\left(n\right)\right]; \\ \Lambda\theta_H\bar{q}^{a^*}\left[A\bar{q}_t + A\eta\bar{q}_t + \bar{q}_t\tilde{B}\left(0\right)\right] \end{array}\right\}. \end{aligned}$$

Once again equating coefficients, we obtain $A = \frac{\pi}{r+\tau}$ and

$$\begin{aligned} \left(r-g+\tau\right)\tilde{B}\left(n\right) &= \psi\Lambda\theta_H\left[A\left(1+\eta\right) + \tilde{B}\left(0\right)\right] \\ &\quad + \max\left\{\begin{array}{c} \xi\left[A\eta\alpha^{n+1} + \tilde{B}\left(n+1\right) - \tilde{B}\left(n\right)\right]; \\ \Lambda\theta_H\bar{q}^{a^*}\left[\left(1+\eta\right)A + \tilde{B}\left(0\right)\right] \end{array}\right\}. \end{aligned} \tag{21}$$

Let us next define $\hat{B}\left(n\right)$ as the solution to the equation:

$$\left(r-g+\tau\right)\hat{B}\left(n\right) = \psi\Lambda\theta_H\left[A\left(1+\eta\right) + \tilde{B}\left(0\right)\right] + \xi\left[\tilde{A}\eta\alpha^{n+1} + \hat{B}\left(n+1\right) - \hat{B}\left(n\right)\right].$$

Under the hypothetical scenario where the max operator in (21) always picks the first term, we have $\tilde{B}(n) = \hat{B}(n)$. Collecting terms, we can write

$$\hat{B}(n) = \beta\hat{\psi}/\xi + \beta\tilde{A}\eta\alpha^{n+1} + \beta\hat{B}(n+1)$$

where $\beta = \frac{\xi}{(r-g+\tau+\xi)}$ and $\hat{\psi} = \psi\Lambda\theta_H\left[A(1+\eta) + \tilde{B}(0)\right]$. From standard dynamic programming arguments (e.g., Theorem 4.7 in Stokey and Lucas, 1989), $\hat{B}(n)$ is strictly decreasing and limits to $\frac{\hat{\psi}}{r-g+\tau}$. Now note that if $n^* = \infty$ (meaning that incremental innovations were always optimal), then we would have $\tilde{B}(n) = \hat{B}(n)$.

The other option in the max operator, $\Lambda\theta_H\bar{q}^{a^*}\left[(1+\eta)A + \tilde{B}(0)\right]$, does not depend on n and is strictly positive, which implies that switching to radical innovation for n sufficiently high would yield $\tilde{B}(n) = \frac{\hat{\psi}+\Lambda\theta_H\bar{q}^{a^*}\left[(1+\eta)A+\tilde{B}(0)\right]}{r-g+\tau} > \frac{\hat{\psi}}{r-g+\tau}$. Hence, there must exist n^* such that firms with $n < n^*$ undertake incremental innovation and switch to radical innovation at n^*. The expression for n^* follows by equating the value of pursuing radical and incremental innovations at n' and setting n^* as the smallest integer greater than n'. ■

Proof of Proposition 5. Recall that

$$\begin{aligned}(r+\tau)V_H(q_{n,t},n) - \dot{V}_H(q_{n,t},n) \quad &= \quad \pi q_{n,t} + \max\left\{\begin{array}{c}\xi\left[V_H\left(q_{n,t}+\eta_{n+1,t}(q_{n,t}),n+1\right) - V_H(q_{n,t},n)\right]; \\ \Lambda\theta_H\bar{q}^a\mathbb{E}V_H(t)\end{array}\right\} \\ &\quad +\psi\Lambda\theta_H\mathbb{E}V_H(t),\end{aligned}$$

where we have written explicitly $\eta_{n+1,t}(q_{n,t})$ as the incremental improvement in productivity starting from quality $q_{n,t}$ that has been improved n times already and average quality in the economy is $\bar{q}_t$ (subsumed in the time argument t).

The threshold number of incremental innovations as a function of current productivity, $n_t^*(q)$ equivalently defines a threshold value of productivity $q_{n,t}^*$ as a function of the number of incremental innovations. Clearly this threshold productivity level is defined as the value that sets the two terms in the max operator equal to each other. Thus

$$V_H\left(q_{n,t}^* + \eta_{n+1,t}(q_{n,t}^*), n+1\right) - V_H\left(q_{n,t}^*, n\right) = \frac{\Lambda\theta_H\bar{q}^a}{\xi}\mathbb{E}V_H(t), \tag{22}$$

and at this value, we also have

$$(r+\tau)V_H\left(q_{n,t}^*, n\right) - \dot{V}_H\left(q_{n,t}^*, n\right) = \pi q_{n,t}^* + \Lambda\theta_H\bar{q}^a\mathbb{E}V_H(t) + \psi\Lambda\theta_H\mathbb{E}V_H(t). \tag{23}$$

Now we will consider two alternative cases:

Case 1:

$$q_{n+1,t}^* \geq q_{n,t}^* + \eta_{n+1,t}(q_{n,t}^*). \tag{24}$$

This condition implies that if a particular high-type firm finds it optimal to switch to radical innovation today, but instead undertakes a successful incremental innovation (as a deviation off-the-equilibrium path), then subsequently it will still want to immediately switch to radical innovation.

Under this case, we have

$$\begin{aligned}
& (r+\tau)V_H\left(q^*_{n,t}+\eta_{n+1,t}(q^*_{n,t}),n+1\right)-\dot{V}_H\left(q^*_{n,t}+\eta_{n+1,t}(q^*_{n,t}),n+1\right) \\
= \;& \pi q^*_{n,t}+\pi\eta_{n+1,t}(q^*_{n,t})+\Lambda\theta_H\bar{q}^a\mathbb{E}V_H(t)+\psi\Lambda\theta_H\mathbb{E}V_H(t).
\end{aligned} \tag{25}$$

This follows from the fact that, by definition, in this case, at $q^*_{n,t}+\eta_{n+1,t}(q^*_{n,t})$, the firm will want to switch to radical innovation.

Now differentiating (22) with respect to time, we have

$$\begin{aligned}
\dot{V}_H\left(q^*_{n,t}+\eta_{n+1,t}(q^*_{n,t}),n+1\right)-\dot{V}_H\left(q^*_{n,t},n\right) &= \frac{\Lambda\theta_H\bar{q}^a}{\xi}\partial\mathbb{E}V_H(t)/\partial t \\
&= \frac{\Lambda\theta_H\bar{q}^a}{\xi}g\mathbb{E}V_H(t),
\end{aligned} \tag{26}$$

where, in the second line, we have used the fact that in a stationary equilibrium $\mathbb{E}V_H(t)$ grows at the rate g. Subtracting (23) from (25) and using (26), we obtain:

$$(r+\tau)[V_H\left(q^*_{n,t}+\eta_{n+1,t}(q^*_{n,t}),n+1\right)-V_H\left(q^*_{n,t},n\right)]=\pi\eta_{n+1,t}(q^*_{n,t})+\frac{\Lambda\theta_H\bar{q}^a}{\xi}g\mathbb{E}V_H(t). \tag{27}$$

Then, combining (22) and (27) we can derive

$$\pi\eta_{n+1,t}(q^*_{n,t})=\frac{r-g+\tau}{\xi}\Lambda\theta_H\bar{q}^a\mathbb{E}V_H(t). \tag{28}$$

In this case, for all q less than $q^*_{n,t}$, it is optimal to switch to radical innovation.

Now let us define

$$v_t\equiv\frac{r-g+\tau}{\pi\xi}\Lambda\theta_H\bar{q}^a\mathbb{E}V_H(t), \tag{29}$$

which is independent of both q and n. Using (29) equation (28) can be written as

$$[\kappa\bar{q}_t+(1-\kappa)q^*_{n,t}]\eta\alpha^{n+1}=v_t, \tag{30}$$

or

$$q^*_{n,t}=\frac{v_t/\eta\alpha^{n+1}-\kappa\bar{q}_t}{1-\kappa}. \tag{31}$$

This equation immediately implies that $q^*_{n,t}$ is increasing in n or equivalently that $n^*_t(q)$ is increasing in q.

We next derive the condition under which (24) indeed applies. For this reason, note that from (30) written for $n+2$ incremental innovations, we have

$$q^*_{n+1,t}=\frac{v_t/\eta\alpha^{n+2}-\kappa\bar{q}_t}{1-\kappa}. \tag{32}$$

Combining equations (31) and (32), we obtain that (24) is satisfied if

$$(1-\kappa)\eta\alpha^{n+2}+\alpha\leq 1. \tag{33}$$

Thus whenever (33) holds (and we are in Case 1), we have the desired result that $n_t^*(q)$ is increasing in q. We next establish that whenever the converse of (33) holds, the same result applies.

Case 2:

$$q_{n+1,t}^* - \eta_{n+1,t}(q_{n,t}^*) < q_{n,t}^*. \tag{34}$$

This implies that if a high-type firm is indifferent between radical and incremental innovation at $n+1^{st}$ prior incremental innovations at time t, then it would have preferred to switch to radical innovation at n^{th} prior incremental innovations. This condition is clearly the complement of (24).

In this case, start with $q_{n+1,t}^*$, which satisfies (25). Under condition (34), $q_{n,t}^*$ satisfies (23), so we again arrive at (22), (28) and (31). But then from (31) $q_{n,t}^*$ is increasing in n or $n_t^*(q)$ is increasing in q.

We next verify that Case 2 applies for the complement of the parameter values for which (33) holds. Note that the same expressions for $q_{n+1,t}^*$ as in (32) again applies under Case 2. Thus the condition for (34) to be satisfied, with an identical argument, is

$$(1-\kappa)\eta\alpha^{n+2} + \alpha > 1,$$

which is indeed the complement of (33).

Consequently, regardless of whether (33) or its converse holds, equation (31) applies, and $q_{n,t}^*$ is increasing in n (or equivalently, $n_t^*(q)$ is increasing in q). This completes the proof of the proposition.

■

References

Abrams, D. S., U. Akcigit, and J. Popadak (2013). Patent Value and Citations: Creative Destruction or Defensive Disruption? National Bureau of Economic Research Working Paper 19763.

Acemoglu, D., U. Akcigit, N. Bloom, and W. R. Kerr (2013). Innovation, Reallocation and Growth. National Bureau of Economic Research Working Paper 18993.

Adorno, T. W., E. Frenkel-Brunswik, D. J. Levinson, and R. N. Sanford (1950). *The Authoritarian Personality.* Harpers.

Akcigit, U. and W. R. Kerr (2010). Growth through Heterogeneous Innovations. National Bureau of Economic Research Working Paper 16443.

Akcigit, U. and W. R. Kerr (2015). Growth through Heterogeneous Innovations, Second Version. Penn Institute for Economic Research Working Paper 15-020.

Alesina, A., P. Giuliano, and N. Nunn (2011). Fertility and the Plough. *American Economic Review,* 101 (3), 499-503.

Arellano, M. and S. Bond (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58(2), 277-297.

Arrow, K. (1962). Economic Welfare and the Allocation of Resources for Invention. *NBER Chapters*, 609-626.

Azoulay, P., G. Manso and J. G. Zivin (2011). Incentives and Creativity: Evidence from the Academic Life Sciences. *RAND Journal of Economics* 42(3), 527-544.

Azoulay, P., J. G. Zivin, and J. Wang (2010). Superstar Extinction. *Quarterly Journal of Economics*. 125(2), 549-589.

Barker, V. L., and G. C. Mueller (2002). CEO Characteristics and Firm R&D Spending. *Management Science,* 48(6), 782-801.

Barro, R. and J. W. Lee (2013). A New Data Set of Educational Attainment in the World, 1950-2010. *Journal of Development Economics*, 104, 184–198.

Benmelech, E., and C. Frydman (2014). Military CEOs. National Bureau of Economic Research Working Paper 19782.

Bennedsen, M., F. Pérez-González, and D. Wolfenzon (2008). Do CEOs Matter? Stanford GSB Working Paper.

Bertrand, M. and A. Schoar (2003). Managing with Style: The Effect of Managers on Firm Policies. *Quarterly Journal of Economics,* 118(4), 1169-1208.

Bloom, N., and J. Van Reenen (2007). Measuring and Explaining Management Practices across Firms and Countries. *Quarterly Journal of Economics,* 122(4), 1351-1408.

Bloom, N., and J. Van Reenen (2010). Why Do Management Practices Differ across Firms and Countries? *Journal of Economic Perspectives,* 24(1), 203-224.

Christensen, C. (1997). *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business Press.

Cyert, R. M. and G. J. March (1963). *A Behavioral Theory of the Firm.* Englewood Cliffs, NJ: Prentice-Hall.

Durkheim, E. (1933). *The Division of Labor*. New York: Macmillan.

Fernandez, R. and A. Fogli (2009). Culture: An Empirical Investigation of Beliefs, Work and Fertility. *American Economic Journal: Macroeconomics*, 1 (1), 146-177.

Fogli A. and L. Veldkamp (2013). Germs, Social Networks and Growth. National Bureau of Economic Research Working Paper 18470.

Foster, L., J. Haltiwanger, and C. J. Krizan (2001). Aggregate Productivity Growth: Lessons from Microeconomic Evidence. *New Directions in Productivity Analysis* (eds. E. Dean, M. Harper, and C. Hulten), University of Chicago Press, 303-372.

Galenson D. and B. Weinberg (1999). Age and the Quality of Work: The Case of Modern American Painters. National Bureau of Economic Research Working Paper 7122.

Galenson D. and B. Weinberg (2001). Creating Modern Art: The Changing Careers of Painters in France From Impressionism to Cubism. *American Economic Review*, 91 (4), 1063-1071.

Garicano, L. (2000). Hierarchies and the Organization of Knowledge in Production. *Journal of Political Economy,* 108(5), 874-904.

Garicano, L., and E. Rossi-Hansberg (2004). Inequality and the Organization of Knowledge. *American Economic Review,* 94(2), 197-202.

Gennaioli, N. and R. La Porta, F. Lopez-de-Silanes, and A. Shleifer (2013). Human Capital and Regional Development, *Quarterly Journal of Economics*, 128(1), 105–164.

Gorodnichenko Y. and G. Roland (2012). Culture, Institutions and the Wealth of Nations. UC Berkeley Working Paper.

Greif, A. (1994). Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies. *Journal of Political Economy*, 102(5), 912-950.

Guiso, L., P. Sapienza, and L. Zingales (2010). Civic Capital as the Missing Link. National Bureau of Economic Research Working Paper 15845.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. National Bureau of Economic Research Working Paper 8498.

Harhoff, D., F. Narin, F. M. Scherer, and K. Vopel (1999). Citation Frequency and the Value of Patented Inventions. *Review of Economics and Statistics*, 81(3), 511-515.

Hofstede, G., G. J. Hofstede, and M. Minkov (2010). *Cultures and Organizations: Software for the Mind*. McGraw-Hill.

Hofstede, G. H. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations*. Sage.

Hurst, E. and B. Pugsley (2011). What Do Small Businesses Do? *Brookings Papers on Economic Activity*, Fall 2011.

Jones, B. F. (2009). The Burden of Knowledge and the Death of the Renaissance Man: Is Innovation Getting Harder? *Review of Economic Studies,* 76 (1), 283-317.

Jones, B. F. (2010). Age and Great Invention. *Review of Economics and Statistics,* 92 (1), 1-14.

Jones, B. F. and B. Weinberg (2011). A Dynamics and Scientific Creativity. *Proceedings of the National Academy of Sciences*, 108 (47), 18910-18914.

Kaplan, S. N., M. M. Klebanov, and M. Sorensen (2012). Which CEO Characteristics and Abilities Matter? *Journal of Finance,* 67(3), 973-1007.

Kirzner, I. M. (1997). Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach. *Journal of Economic Literature,* 35 (1), 60-85.

Klette, T. J. and S. Kortum (2004). Innovating Firms and Aggregate Innovation. *Journal of Political Economy*, 112, 986-1018.

Kluckhohn, F. R. and F. L. Strodtbeck (1961). *Variations in Value Orientations.* Evanston, IL: Row Peterson.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2012). Technological Innovation, Resource Allocation, and Growth. National Bureau of Economic Research Working Paper 17769.

Lai, R., A. D'Amour, A. Yu, Y. Sun, and L. Fleming (2009). Disambiguation and Co-author Networks of the US Patent Inventor Database. Harvard Business School Working Paper.

Lentz, R., and D. Mortensen (2008). An Empirical Model of Growth through Product Innovation. *Econometrica,* 76 (2008), 1317-1373.

Lucas Jr, R. (1978). On the Size Distribution of Business Firms. *Bell Journal of Economics,* 9(2), 508-523.

MacDonald, G., and M. S. Weisbach (2004). The Economics of Has-beens. *Journal of Political Economy,* 112(1), S289-S310.

McClelland, D. (1961). *The Achieving Society*. The Free Press, New York.

Parsons, T. (1949). *Essays in Sociological Theory; Pure and Applied*. Free Press.

Sampat, B. N. and A. A. Ziedonis (2004). Patent Citations and the Economic Value of Patents. In *Handbook of Quantitative Science and Technology Research*, 277-298. Springer.

Sarada, and Tocoian , O. (2013). The Success of Entrepreneurial Networks: Evidence from Brazil. Duke University Working Paper.

Schumpeter, J. A. (1934). The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship.*

Schwartz, S. H. (1994). *Beyond Individualism/Collectivism: New Cultural Dimensions of Values.* Sage Publications, Inc.

Schwartz, S. H. and W. Bilsky (1990). Toward a Theory of the Universal Content and Structure of Values: Extensions and Cross-cultural Replications. *Journal of Personality and Social Psychology,* 58(5), 878.

Shane, S. (1993). Cultural Influences on National Rates of Innovation. *Journal of Business Venturing,* 8(1), 59-73.

Shane, S. (1995). Is the Independent Entrepreneurial Firm a Valuable Organizational Form? In *Academy of Management Proceedings*, 110-113. Academy of Management.

Stokey, N. and R. Lucas Jr. (1989) *Recursive Methods in Economic Dynamics*, Harvard University Press.

Syverson, C (2011). What Determines Productivity? *Journal of Economic Literature*, 49(2), 326–65.

Tabellini, G. (2008a). Institutions and Culture. *Journal of European Economic Association*, 6(2), 255-294.

Tabellini, G. (2008b). The Scope of Cooperation: Values and Incentives. *Quarterly Journal of Economics*, 123(3), 905-950.

Trajtenberg, M. (1990). A Penny for Your Quotes: Patent Citations and the Value of Innovations. *RAND Journal of Economics,* 172-187.

Triandis, H. C. (1995). *Individualism & Collectivism.* Westview Press.

Walter, F., and S. Scheibe (2013). A Literature Review and Emotion-based Model of Age and Leadership: New Directions for the Trait Approach. *Leadership Quarterly,* 24(6), 882-901.

Weinberg B. and D. Galenson (2005). Creative Careers: Lifecycles of Noble Laureates in Economics. National Bureau of Economic Research Working Paper 11799.

Weitzman, M. (1998). Recombinant Growth. *Quarterly Journal of Economics,* 113(2), 331-360.

Winslow, E. K. and G. T. Solomon (1987). Entrepreneurs Are More Than Non-Conformists: They Are Mildly Sociopathic. *Journal of Creative Behavior,* 21(3), 202-213.

Table 1: Summary Statistics

Panel A: Descriptive Statistics

Variable	Observations	Mean	Standard Deviation
Balanced Firm Sample (Firm Averages, 1995-2000)			
CEO age	279	55.3	6.47
average manager age	279	52.3	4.32
innovation quality	279	20.5	8.76
superstar fraction	279	12.3	10.1
tail innovation	279	2.72	2.56
generality	279	21.5	5.53
log patents	279	5.86	1.51
log employment	279	3.84	1.38
log sale	279	4.34	1.47
firm age	279	37.3	14.4
R&D intensity	257	8.52	17.0
Unbalanced Firm Sample (Annual Firm Observations, 1992-2004)			
CEO age	7111	55.3	6.84
average manager age	7111	52.3	4.38
innovation quality	7111	15.9	10.9
superstar fraction	7111	9.91	10.7
tail innovation	5803	3.41	5.42
generality	6232	18.5	9.96
log patents	7111	5.61	1.60
log employment	7111	3.71	1.51
log sale	7111	4.12	1.61
firm age	7111	35.1	16.3
Cross-Country Sample (Country Averages, 1995-2000)			
individualism	50	.813	.263
uncertainty aversion	50	.492	.195
average manager age	37	56.1	2.98
innovation quality	50	14.5	3.26
superstar fraction	50	6.68	3.65
tail innovation	50	1.92	.945
generality	50	21.0	1.81
log patents	50	10.5	1.52
log income per capita	50	10.3	.305
secondary years of schooling	50	4.84	.827
R&D intensity	44	2.59	.363

- Table 1 continued on next page -

Panel B: Correlation Matrix of Firm-Level Innovation Variables

	innovation quality	superstar fraction	tail innovation	generality
innovation quality	1.000			
superstar fraction	0.925	1.000		
tail innovation	0.893	0.829	1.000	
generality	-0.177	-0.204	-0.145	1.000

Panel C: Correlation Matrix of Cross-Country Innovation Variables

	innovation quality	superstar fraction	tail innovation	generality
innovation quality	1.000			
superstar fraction	0.932	1.000		
tail innovation	0.945	0.990	1.000	
generality	0.902	0.880	0.906	1.000

Panel D: Correlation Matrix of Openness to Disruption Variables

	individualism	uncertainty aversion	average manager age
individualism	1.000		
uncertainty aversion	-0.884	1.000	
average manager age	-0.770	0.844	1.000

Notes: All statistics in this table are weighted by the number of patents (of the country or the firm). Individualism and uncertainty aversion are Hofstede's indices of national cultures (and are normalized to lie between 0 and 1), and country average manager age is the average manager of CEOs and CFOs of up to the 25 largest firms in the country. Innovation quality is the average number of citations per patent (using the truncation correction weights devised by Hall, Jaffe, and Trajtenberg, 2001); superstar fraction is the fraction of patents accounted for by superstar researchers (those above the 95th percentile of the citation distribution); tail innovation is the fraction of patents of a country or firm above the 99th percentile of the citation distribution divided by the fraction of patents above the median of the distribution; and generality index measures the dispersion of citations received across two-digit IPC technology classes. Log income per capita at the country level, and log employment, log sales at the firm level are computed as the average of, respectively, annual log income per capita, log employment and log sale between 1995 and 2000. CEO age is the age of the CEO and average manager age is the average age of the top management, both from the Execucomp dataset. The balanced firm panel is the sample of firms from Compustat with complete data on CEO age, employment, sales, and firm age and positive patents in each year between 1995 and 2000. The unbalanced firm panel is a sample of firms from Compustat with at least one year of complete data between 1992 and 2002. See text for the definition of other variables and further details.

Table 2: Baseline Firm-Level Regressions

	Innovation Quality	Superstar Fraction	Tail Innovation	Generality
CEO age	-0.278	-0.300	-0.151	-0.183
	(0.088)	(0.141)	(0.054)	(0.055)
firm age	-0.219	-0.238	-0.063	0.029
	(0.078)	(0.106)	(0.029)	(0.046)
log employment	-1.599	-4.813	-0.908	-4.574
	(1.937)	(3.376)	(0.793)	(1.500)
log sales	1.833	5.215	0.743	4.421
	(1.425)	(2.645)	(0.650)	(1.331)
log patent	1.073	0.093	0.662	-0.696
	(0.769)	(1.336)	(0.356)	(0.633)
R^2	0.88	0.81	0.79	0.83
N	279	279	279	279

Notes: Weighted cross-sectional regressions with total number of patents as weights. The sample is the balanced firm panel and each observation is the sample average between 1995-2000 as described in the notes to Table 1. The dependent variables are innovation quality, superstar fraction, tail innovation, and generality (the last three are multiplied by 100 to ease legibility). In addition, all regressions control for a full set of dummies for four-digit SIC industries. See text and notes to Table 1 for variable definitions. Robust standard errors are in parentheses.

Table 3: Firm-Level Regressions (Alternative Measures)

	Innovation Quality (5 years)	Superstar Fraction (Best Patent)	Tail Innovation (90/50)	Originality
CEO age	-0.129	-0.497	-0.299	-0.285
	(0.041)	(0.332)	(0.094)	(0.075)
R^2	0.87	0.87	0.83	0.87
N	279	279	279	279

Notes: Weighted cross-sectional regressions with total number of patents as weights. The sample is the balanced firm panel and each observation is the sample average between 1995-2000 as described in the notes to Table 1. The dependent variables are alternative measures of innovation quality (computed over the next five years), superstar fraction (with superstars defined according to the best patent), tail innovation (with share of the patents of the firm among all the patents above the 90th percentile of the citation distribution in the numerator), and the originality index (the last three are multiplied by 100 to ease legibility). All regressions control for firm age, log employment, log sales, log total patents, and a full set of dummies for four-digit SIC industries. See text and notes to Table 1 for variable definitions. Robust standard errors are in parentheses.

Table 4: Firm-Level Regressions (Robustness)

	Innovation Quality	Superstar Fraction	Tail Innovation	Generality
Panel A: With SIC3 Dummies				
CEO age	-0.257	-0.284	-0.126	-0.086
	(0.070)	(0.123)	(0.050)	(0.091)
R^2	0.77	0.72	0.64	0.70
N	279	279	279	279
Panel B: With Additional Controls				
CEO age	-0.270	-0.282	-0.150	-0.194
	(0.090)	(0.140)	(0.052)	(0.054)
R^2	0.88	0.82	0.79	0.83
N	279	279	279	279
Panel C: With Additional Controls Plus R&D Intensity				
CEO age	-0.258	-0.295	-0.142	-0.184
	(0.088)	(0.149)	(0.048)	(0.053)
R^2	0.89	0.82	0.81	0.84
N	257	257	257	257
Panel D: With Average Manager Age				
average manager age	-0.418	-0.467	-0.224	-0.339
	(0.163)	(0.206)	(0.094)	(0.084)
R^2	0.87	0.81	0.77	0.83
N	279	279	279	279
Panel E: High-Tech Subsample				
CEO age	-0.227	-0.191	-0.145	-0.189
	(0.068)	(0.157)	(0.045)	(0.043)
R^2	0.92	0.84	0.86	0.81
N	87	87	87	87
Panel F: Low-Tech Subsample				
CEO age	-0.439	-0.704	-0.143	-0.153
	(0.200)	(0.252)	(0.085)	(0.146)
R^2	0.85	0.82	0.72	0.86
N	192	192	192	192

Notes: Weighted cross-sectional regressions with total number of patents as weights. The sample is the balanced firm panel and each of the ratios is the sample average 1995-2000 as described in the notes to Table 1.The dependent variables are innovation quality, superstar fraction, tail innovation, and generality (the last three are multiplied by 100 to ease legibility). Each panel is for a different specification. Unless otherwise stated, all regressions control for firm age, log employment, log sales, log total patents, and four-digit SIC dummies (see text and notes to Table 1 for variable definitions). Robust standard errors are in parentheses. Panel A controls for three-digit SIC dummies instead of the four-digit dummies. Panel B adds to the specification of Table 2 profitability (profit over sales), indebtedness (debt over sales) and log physical capital. Panel C adds to the specification of Panel B R&D intensity (R&D expenditure over sales). Panel D uses average manager age instead of CEO age. Panels E and F are for the high-tech and low-tech subsamples. High-tech sample includes all firms with a primary industry classification of SIC 35 (industrial and commercial machinery and equipment and computer equipment) and 36 (electronic and other electrical equipment and components), while the low-tech sample includes the rest.

Table 5: Firm-Level Panel Regressions

	Innovation Quality	Superstar Fraction	Tail Innovation	Generality
Panel A: Average CEO Age (No Fixed Effects), Balanced Firm Sample, 1995-2000				
average CEO age	-0.227	-0.336	-0.132	-0.183
	(0.068)	(0.103)	(0.041)	(0.044)
R^2	0.70	0.69	0.47	0.75
N	1,674	1,674	1,594	1,655
Panel B: Average CEO Age (No Fixed Effects), Unbalanced Firm Sample, 1992-2004				
average CEO age	-0.168	-0.319	-0.104	-0.171
	(0.075)	(0.133)	(0.045)	(0.044)
R^2	0.66	0.54	0.31	0.77
N	7,111	7,111	5,803	6,232
Panel C: CEO Age (Fixed Effects), Unbalanced Firm Sample, 1992-2004				
CEO age	-0.188	-0.149	-0.076	0.036
	(0.044)	(0.051)	(0.023)	(0.029)
R^2	0.78	0.80	0.44	0.85
N	7,111	7,111	5,803	6,232
Panel D: CEO Age and Lagged CEO Age (Fixed Effects), Unbalanced Firm Sample, 1993-2004				
CEO age	-0.131	-0.098	-0.052	0.031
	(0.041)	(0.039)	(0.023)	(0.026)
lagged CEO age	-0.123	-0.100	-0.055	0.020
	(0.051)	(0.049)	(0.029)	(0.035)
R^2	0.80	0.81	0.46	0.85
N	5,407	5,407	4,562	4,780
Panel E: CEO Age and Lagged Dependent Var (Fixed Effects), Unbalanced Firm Sample, 1993-2004				
CEO age	-0.096	-0.075	-0.065	0.037
	(0.026)	(0.030)	(0.019)	(0.024)
lagged dependent variable	0.472	0.452	0.194	0.200
	(0.034)	(0.046)	(0.051)	(0.042)
R^2	0.86	0.86	0.46	0.86
N	5,985	5,985	4,772	5,207
Panel F: CEO Age and Lead CEO Age (Fixed Effects), Unbalanced Firm Sample, 1992-2003				
CEO age	-0.113	-0.084	-0.042	0.042
	(0.042)	(0.048)	(0.019)	(0.029)
lead CEO age	-0.125	-0.109	-0.043	-0.007
	(0.049)	(0.044)	(0.022)	(0.028)
R^2	0.78	0.81	0.48	0.85
N	5,409	5,409	4,849	5,097

Notes: Weighted firm-level panel regressions with annual observations with number of patents (in that year) as weights. The dependent variables are innovation quality, superstar fraction, tail innovation, and generality (the last three are multiplied by 100 to ease legibility). Robust standard errors clustered at the firm level are in parentheses. Panel A is for our balanced firm sample 1995-2000, and controls for firm age, log employment, log sales, log patents, a full set of four-digit SIC dummies, and year dummies (and thus no firm dummies), and the key right-hand side variable is average CEO age (constant over time). Panel B is identical to Panel A except that the sample is extended to the unbalanced firm panel 1992-2002. In Panel C, the key right-hand side variable is CEO age (in that year), and the regression also includes a full set of firm fixed effects (and thus firm age and the four-digit SIC dummies are no longer included). Panel D is identical to Panel C except that it also includes a one year lag of CEO age as well as current CEO age, and Panel E is identical to Panel C except that it also includes a one year lag of the dependent variable on the right-hand side. See text and notes to Table 1 for variable definitions.

Table 6: Structural Parameters

Parameter	*Description*	*Identification*
	External Calibration	
$x = 0.05$	Entry rate	Compustat sample
$\rho = 0.02$	Discount rate	Standard value
$\delta = 0.04$	Manager death rate	Compustat sample
$\alpha = 0.93$	Reduction rate of innovation size	Akcigit and Kerr (2015)
	Indirect Inference	
$\psi = 10.2$	Baseline radical innovation rate for high type	Estimate
$\Lambda\theta_H = 0.005$	High-type innovation parameter	Estimate
$\varphi = 0.149$	Transition rate from low type to high type	Estimate
$\xi = 0.031$	Incremental innovation rate	Estimate
$\eta = 0.449$	Initial innovation size	Estimate
$\zeta = 0.254$	Probability of high-type entrant	Estimate

Notes: Parameter choices and estimates. See Section 4.4 for details.

Table 7: Empirical and Model-Generated Moments

Target	*U.S. Data*	*Model*
Current manager age coefficient of Table 5 Panel C	−0.188	−0.211
Lead manager age coefficient of Table 5 Panel F	−0.125	−0.129
Current manager age coefficient of Table 5 Panel F	−0.113	−0.111
Annual growth rate	5.75%	5.39%
Within-firm coefficient of variation of radical innovations	1.99	2.17
Fraction of internal patents	21.5%	23.8%

Notes: Empirical and model-generated moments for the indirect inference procedure. See Section 4.4 for details.

Table 8: Patent-Level Panel Regressions

	Innovation Quality	Tail Innovation (Above 99)	Tail Innovation (Above 90)	Generality
Panel A: CEO Age, Unbalanced Firm Sample, 1992-2004				
CEO age	-0.119	-0.314	-1.239	0.028
	(0.038)	(0.132)	(0.413)	(0.025)
R^2	0.11	0.03	0.07	0.11
N	316,516	316,516	316,516	263,641
Panel B: Inventor Age, Unbalanced Firm Sample, 1992-2004				
inventor age	-0.234	-0.440	-2.883	-0.019
	(0.026)	(0.121)	(0.321)	(0.022)
R^2	0.14	0.03	0.09	0.15
N	316,516	316,516	316,516	263,641
Panel C: Inventor Age, Extended Sample, 1985-2004				
inventor age	-0.226	-0.377	-2.842	-0.017
	(0.022)	(0.075)	(0.293)	(0.017)
R^2	0.16	0.05	0.10	0.15
N	572,169	572,169	572,169	466,378
Panel D: Inventor Age, Extended Sample, 1985-2004				
inventor age	-0.201	-0.327	-2.359	-0.046
	(0.010)	(0.036)	(0.134)	(0.011)
R^2	0.27	0.15	0.19	0.25
N	1,855,887	1,855,887	1,855,887	1,550,825
Panel E: CEO Age and Inventor Age, Unbalanced Firm Sample, 1992-2004				
inventor age	-0.233	-0.438	-2.876	-0.019
	(0.026)	(0.121)	(0.321)	(0.022)
CEO age	-0.119	-0.317	-1.218	0.028
	(0.036)	(0.126)	(0.388)	(0.022)
R^2	0.14	0.03	0.09	0.15
N	316,516	316,516	316,516	263,641

Notes: Patent-level panel regressions with annual observations. The dependent variables are innovation quality at the patent level; a dummy for the patent being above the 99th percentile of the citation distribution; dummy for the patent being above the 90th percentile of the citation distribution; and generality index at the patent level (the last three are multiplied by 100 to ease legibility). Robust standard errors clustered at the firm level are in parentheses. Panel A is for our unbalanced firm sample 1992-2002 and controls for log employment, log sales, log patents, a full set of firm fixed effects, and application year dummies, and the key right-and side variable is CEO age. Panel B is for our unbalanced firm sample 1992-2002 and controls for log employment, log sales, log patents, application year dummies, a full set of firm fixed effects, a full set of dummies for inventor team size, a full set of dummies for three-digit IPC technology class dummies, and a full set of dummies for the total number of patents of the inventor within the team with the highest number of patents, and the key right-and side variable is average inventor age. Panel C expands the sample of Panel B to 1985-2002 and also adds Compustat firms without CEO information into the sample. Panel D extends the sample of Panel C to include non-Compustat firms as well (hence excludes log sales and log employment, and still includes a full set of firm fixed effects). Panel E is for our unbalanced firm sample 1992-2002 and adds CEO age to the specification of Panel B. See text and notes to Table 1 for variable definitions.

Table 9: Inventor Age and CEO Age,
Unbalanced Firm Sample, 1992-2004

	Inventor age	Inventor age
	(1)	(2)
CEO age	0.014	0.013
	(0.006)	(0.002)
R^2	0.11	0.13
N	316,516	316,516

Notes: Patent-level panel regressions with annual observations for the unbalanced firm sample 1992-2002. The dependent variable is the average age of inventors. The first column controls for log employment, log sales, log patents, application year dummies, and a full set of firm fixed effects, and the second column adds to this a full set of team size dummies and a full set of dummies for three-digit IPC technology class dummies. See text and notes to Table 1 for variable definitions.

Table 10: Stock of Knowledge, Opportunity Cost, and Creative Innovations,
Unbalanced Firm Sample, 1992-2004

	Innovation Quality	Superstar Fraction	Tail Innovation	Generality
CEO age	-0.180	-0.216	-0.087	-0.044
	(0.027)	(0.027)	(0.017)	(0.016)
log sales	1.465	2.081	0.285	1.201
	(0.449)	(0.611)	(0.272)	(0.328)
log patent	-0.394	-0.072	0.391	-0.020
	(0.193)	(0.257)	(0.136)	(0.151)
CEO age $\times$ log patent	-0.005	-0.071	-0.016	-0.037
	(0.014)	(0.021)	(0.011)	(0.011)
CEO age $\times$ log sales	0.024	0.079	0.009	0.044
	(0.017)	(0.021)	(0.012)	(0.011)
R^2	0.67	0.55	0.31	0.77
N	7,111	7,111	5,803	6,232

Notes: Weighted firm-level panel regressions with annual observations for the unbalanced firm panel, 1992-2002, with number of patents (in that year) as weights. The dependent variables are innovation quality, superstar fraction, tail innovation, and generality (the last three are multiplied by 100 to ease legibility). Robust standard errors clustered at the firm level are in parentheses. All regressions also include log employment, application year dummies and a full set of dummies for four-digit SIC industries. See text and notes to Table 1 for variable definitions.

Table 11: Baseline Cross-Country Regressions

	Innovation Quality	Superstar Fraction	Tail Innovation	Generality
	Panel A: Average Manager Age			
manager age	-0.484	-0.960	-0.225	-0.278
	(0.225)	(0.221)	(0.058)	(0.056)
log income per capita	-0.491	-0.702	-0.136	0.211
	(1.153)	(1.066)	(0.291)	(0.468)
secondary years of schooling	-1.000	-1.359	-0.291	-0.231
	(1.481)	(1.462)	(0.396)	(0.341)
log patent	2.232	2.331	0.591	1.072
	(0.706)	(0.695)	(0.193)	(0.222)
R^2	0.74	0.82	0.80	0.80
N	37	37	37	37
	Panel B: Individualism			
individualism	4.965	9.929	2.369	3.420
	(2.461)	(2.393)	(0.640)	(0.487)
log income per capita	-1.233	-2.130	-0.472	-0.252
	(1.195)	(1.270)	(0.334)	(0.373)
secondary years of schooling	-0.467	-0.317	-0.056	-0.051
	(1.229)	(1.174)	(0.323)	(0.227)
log patents	1.622	1.125	0.308	0.725
	(0.490)	(0.472)	(0.129)	(0.164)
R^2	0.73	0.81	0.79	0.83
N	50	50	50	50
	Panel C: Uncertainty Avoidance			
uncertainty avoidance	-8.354	-13.528	-3.174	-4.242
	(2.946)	(2.715)	(0.722)	(0.798)
log income per capita	-0.408	-0.657	-0.124	0.232
	(0.957)	(0.600)	(0.177)	(0.558)
secondary years of schooling	-0.745	-0.346	-0.054	0.008
	(1.149)	(1.108)	(0.307)	(0.208)
log patent	1.708	1.257	0.339	0.765
	(0.439)	(0.424)	(0.125)	(0.189)
R^2	0.80	0.86	0.84	0.84
N	50	50	50	50

Notes: Weighted cross-country regressions with total number of patents as weights. The dependent variables are innovation quality, superstar fraction, tail innovation, and generality (the last three are multiplied by 100 to ease legibility). See text and notes to Table 1 for variable definitions. Each country observation is the sample average between 1995-2000 as described in the text and the notes to Table 1. Robust standard errors are in parentheses.

Table 12: Cross-Country Regressions (Alternative Measures)

	Innovation Quality (5 years)	Superstar Fraction (Best Patent)	Tail Innovation (90/50)	Originality
Panel A: Average Manager Age				
manager age	-0.203	-0.005	-1.002	-0.713
	(0.092)	(0.004)	(0.372)	(0.083)
R^2	0.75	0.80	0.70	0.88
N	37	37	37	37
Panel B: Individualism				
individualism	2.039	0.052	9.966	8.015
	(1.009)	(0.045)	(4.028)	(0.653)
R^2	0.74	0.80	0.68	0.91
N	50	50	50	50
Panel C: Uncertainty Avoidance				
uncertainty avoidance	-3.461	-0.106	-15.964	-9.084
	(1.215)	(0.057)	(4.689)	(1.336)
R^2	0.81	0.83	0.78	0.87
N	50	50	50	50

Notes: Weighted cross-country regressions with total number of patents as weights. The dependent variables are alternative measures of innovation quality (computed over the next five years), superstar fraction (with superstars defined according to the best patent), tail innovation (with fraction of patents above the 90th percentile of the citation distribution in the numerator), and the originality index (the last three are multiplied by 100 to ease legibility). Each regression also controls for log income per capita, average years of secondary schooling, and log total patents. See text and notes to Table 1 for variable definitions. Each country observation is the sample average between 1995-2000 as described in the text and the notes to Table 1. Robust standard errors are in parentheses.

Table 13: Cross-Country Regressions (Controlling for R&D Intensity)

	Innovation Quality	Superstar Fraction	Tail Innovation	Generality
Panel A: Average Manager Age				
manager age	-0.636	-1.096	-0.257	-0.622
	(0.255)	(0.253)	(0.066)	(0.105)
R^2	0.76	0.83	0.81	0.91
N	33	33	33	33
Panel B: Individualism				
individualism	8.245	13.786	3.291	2.932
	(2.821)	(2.602)	(0.725)	(0.778)
R^2	0.78	0.85	0.83	0.83
N	44	44	44	44
Panel C: Uncertainty Avoidance				
uncertainty avoidance	-9.589	-14.173	-3.305	-3.452
	(2.747)	(2.753)	(0.754)	(0.915)
R^2	0.82	0.86	0.83	0.85
N	44	44	44	44

Notes: Weighted cross-country regressions with total number of patents as weights. The dependent variables are innovation quality, superstar fraction, tail innovation, and generality (the last three are multiplied by 100 to ease legibility). Each regression also controls for log income per capita, average years of secondary schooling, log total patents, and R&D intensity defined as total R&D expenditure divided by GDP. See text and notes to Table 1 for variable definitions. Each country observation is the sample average between 1995-2000 as described in the text and the notes to Table 1. Robust standard errors are in parentheses.

Table A1: Average Annual Patent Counts by Country, 1995-2000

Country	*Abbreviation*	*Patent Count*	*Country*	*Abbreviation*	*Patent Count*
Argentina	AR	9.2	India	IN	90.3
Austria	AT	365.0	Italy	IT	1439.8
Australia	AU	744.0	Japan	JP	33954.8
Belgium	BE	522.8	South Korea	KR	3581.5
Bulgaria	BG	3.8	Luxemburg	LU	62.8
Brazil	BR	69.7	Malta	MT	2.0
Canada	CA	2433.2	Mexico	MX	59.2
Switzerland	CH	1588.7	Malaysia	MY	14.5
Chile	CL	8.8	Netherlands	NL	1236.7
China	CN	109.5	Norway	NO	239.2
Colombia	CO	2.0	New Zealand	NZ	104.7
Czech Republic	CZ	17.7	Poland	PL	10.0
Germany	DE	9257.0	Portugal	PT	8.7
Denmark	DK	448.5	Romania	RO	2.7
Spain	ES	193.8	Russia	RU	88.2
Finland	FI	910.3	Saudi Arabia	SA	18.2
France	FR	3877.5	Sweden	SE	1691.3
Great Britain	GB	2869.5	Singapore	SG	191.2
Greece	GR	15.7	Slovenia	SI	13.7
Hong Kong	HK	171.8	Slovakia	SK	4.0
Croatia	HR	7.7	Thailand	TH	10.7
Hungary	HU	33.3	Turkey	TR	5.3
Indonesia	ID	3.0	United States	US	93722.5
Ireland	IE	111.3	Venezuela	VE	24.3
Israel	IL	580.7	South Africa	ZA	88.7

Notes: This table shows the average annual patent counts between 1995-2000, registered at the USPTO from that country.

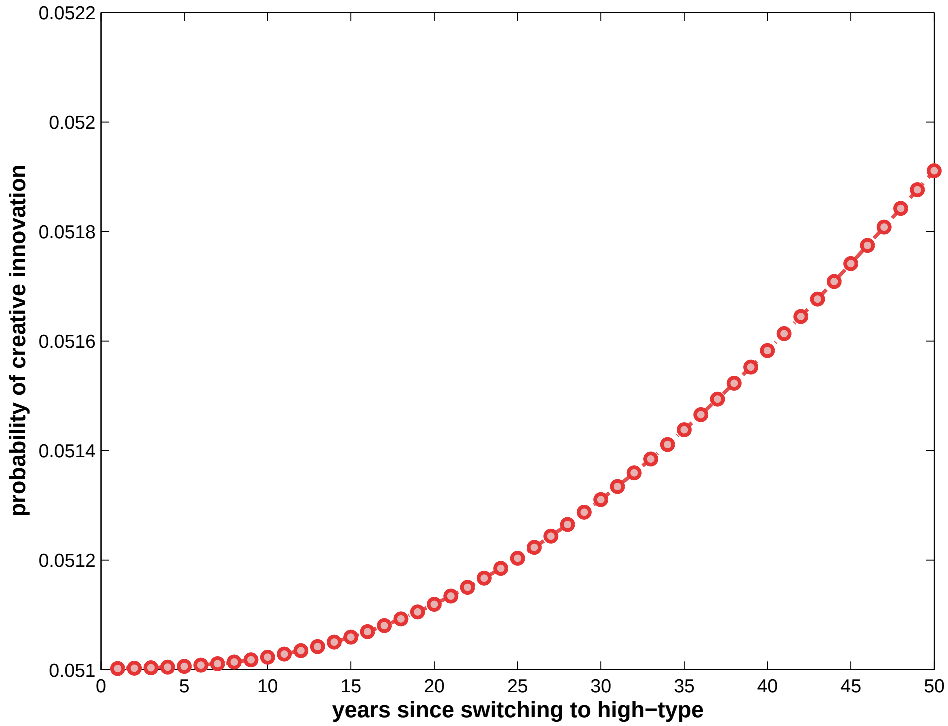

Figure 1: Evolution of creative innovations for high-type firms.

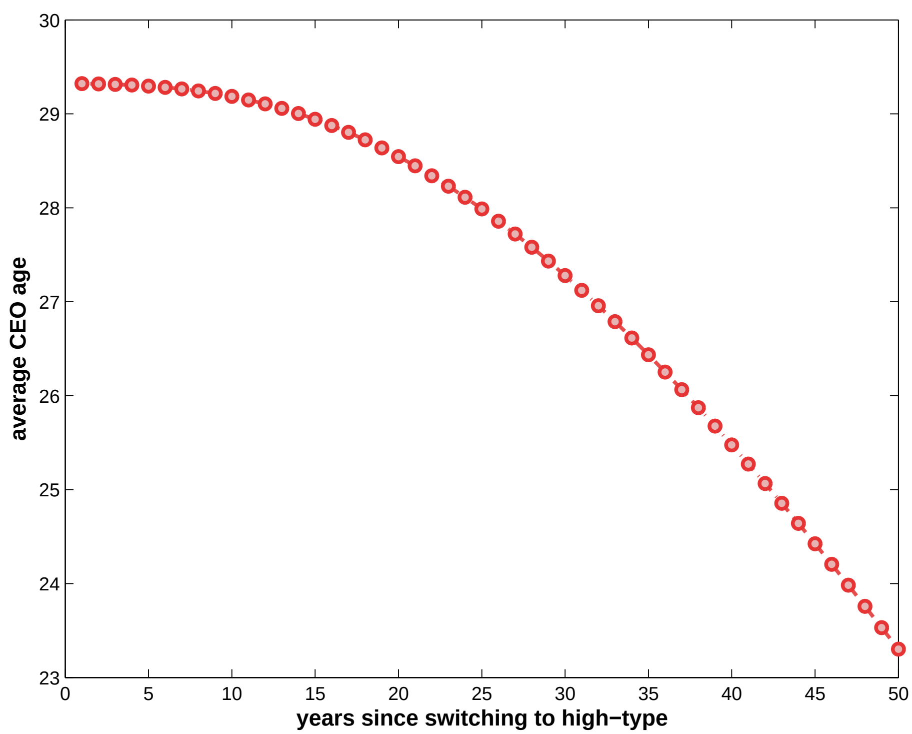

Figure 2: Evolution of CEO age for high-type firms.