NBER WORKING PAPER SERIES

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

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 Pascual Restrepo and conference and seminar participants at the NBER Productivity group, CIFAR, Brown University, University of Pennsylvania, 2014 AEA Meetings and particularly our discussant Joshua Gans, for helpful suggestions and comments. We also thank the Bilkent University Economics Department for the great hospitality during this project. Financial support from 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 Innovationsˆ
Daron Acemoglu, Ufuk Akcigit, and Murat Alp Celik
NBER Working Paper No. 19894
February 2014
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 types of firms, we provide cross-country, firm-level and patent-level 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 manager age. This variable is 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 country, firm or patent level, we present robust evidence that openness to disruption is associated with more creative innovations.

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 Pennsylvania
3718 Locust Walk, #445
Philadelphia, PA 19104
and NBER
uakcigit@econ.upenn.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 per year granted by the US Patent and Trademark Office (USPTO), only a handful of those 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 (e.g., Hall, Jaffe and Trajtenberg, 2001, and further references discussed below). 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 comes from 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 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 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).

At the cross-country level, these expectations are borne out when looking at the relationship between various measures of creative innovations and several proxies for openness to disruption. Figure 1 gives a glimpse of the cross-country patterns, which we document further below. This figure uses the *average number of citations per patent* filed with the USPTO originating from different countries between the years 1995 and 2000 as the dependent variable in all three panels (see Section 3 for more information on data and variables). Because of the importance of the US market for international businesses, companies located all around the world file their more important innovations with the USPTO, giving us a decent sample size of patents from 50 countries (though with some notable differences in the number of patents). The average number of citations per patent is a proxy for the average quality of innovations, since higher-quality patents tend to get cited more (below we use several other measures as alternative proxies, which all show similar patterns). Each panel of Figure 1 uses a different proxy for openness to disruption and depicts the conditional relationship between the average number of citations per patent and the proxy in question (after controlling for log GDP per capita, average years of secondary schooling and log total number of patents in the country). It also shows the weighted regression line.[1]

The first panel focuses on the "individualism" variable of the Dutch social scientist Geert Hofstede, who constructed indices for various dimensions of "national cultures." This variable, based on Durkheim's (1933) distinction between collectivism and individualism, 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.[2] The second panel uses Hofstede's "uncertainty avoidance" index, which 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.

The third panel uses our own proxy for a society's openness to disruption: the average age of (top) managers (e.g., CEO and CFO) in the 25 largest listed companies in the country (when available). The motivation for this variable is that societies that are open to disruption tend to be more meritocratic in promoting young talent, even those promoting and implementing disruptive innovations, while those that discourage individualistic risk-taking attitudes tend to make the young subservient to the old.[3]

[1] This is from a weighted regression the using total number of patents as weights to partially correct for the fact that the number of patents is small and thus our measures are quite noisy for several countries in our sample.

[2] All of our variables are defined and described in greater detail below.

[3] 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).

The average manager age variable, which we collected from publicly available sources, has several advantages: as opposed to Hofstede's data it is not subjective; it closely corresponds to a specific dimension of openness to disruption, and it is available for US companies, enabling us to conduct most of our analysis with firm-level and patent-level data within the United States.

In all three panels, there is a fairly strong positive relationship, and the weighted regression line is statistically significant at 5% or less (see below). It is worth noting that this relationship does not reflect the potential correlation between our measures of creative innovations and GDP per capita, human capital, or even total number of patents, each of which is controlled for in the plots and the corresponding regressions. Though far from conclusive, this evidence shows a notable pattern in the cross-country data on creative innovations, which, to the best of our knowledge, has neither been noted nor systematically investigated before.

Motivated by these patterns, we first provide a simple model of firm innovation strategies. Firms can engage in an *incremental innovation* by building on their existing leading-edge products or a *radical innovation* by combining diverse ideas to generate an improvement in a new area. We assume that some companies have a comparative advantage in radical innovations (because of their technology, "type" or "corporate culture"), but in addition, their managers' skills are also important for the type of innovation. In particular, young managers have more recently acquired general skills (or are less beholden to a particular type of product or technology). This enables them to more effectively use current advances in a range of fields to succeed in 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 relationship between manager age and radical innovation. But this is not necessarily the causal effect of manager age. Rather, manager age is both an economically relevant variable and more generally a proxy for openness to disruption. In the model, this is captured by the fact that there is both sorting of young managers to firms that are open to radical innovation, and also young managers employed by such firms do contribute to their radical innovations. The model further clarifies that radical innovations will generate higher quality patents that are more likely to receive a very 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). It further 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).

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 the more recent skills of such managers that are necessary for 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, though the firm-level and patent-level results are our main focus.

The bulk of our paper comprises an empirical investigation of the ideas proposed so far and illustrated by our theoretical model. After describing the datasets and the various measures we use to proxy for creative innovations in Section 3, we provide a few more details on cross-country relationships in Section 4. This involves presenting the regression evidence corresponding to Figure 1 and the results with alternative measures of creative innovations that we also use in our firm-level analysis. In particular, in addition to the *average number of citations per patent*, we use three other main measures: *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. Our results in Section 4 show that patterns similar to those shown in Figure 1 are also present with these three additional quite distinct measures of creative innovations.

Section 5 then turns to our main empirical focus: the firm-level and patent-level analysis of openness to disruption and creative innovations. We work with the Compustat sample and use the age of the CEO (or the average age of top management for robustness) as our proxy for openness to disruption. We find a very robust correlation between this proxy of openness to disruption and all of our measures of firm-level creative innovation (with or without a variety of firm-level controls). Even though we have no compelling strategy to identify an exogenous source of variation in CEO age (or in openness to disruption at the company level), the firm-level correlations we present are quite robust.

Perhaps surprisingly, we find similar results when we control for firm fixed effects and exploiting within firm variation, so that when a younger CEO takes charge, innovations (new patent applications) become more creative. Exploiting patent-level variation, we also estimate the impact of

CEO and inventor age on the creativity of innovations. Our results indicate that both matter, with roughly similar magnitudes. Interestingly, 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).

We further 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 average CEO age with (log) sales and (log) number of patents of the firm. Though the results here are a little 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.

Our paper is related to several literatures. 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." 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).

Most closely related to our work are recent papers by Gorodnichenko and Roland (2012) and Fogli and Veldkamp (2012). Gorodnichenko and Roland also 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 from 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. Though we also show results with Hofstede's data, this is only to provide motivating evidence. 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 the individualism index in their theoretical and empirical analysis of "individualistic" social 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.

Also closely linked is 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 on, but they also acquire other types of human capital (perhaps generating different types of creativity) later in their careers. 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.[4]

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 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 recent 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.

Our paper also relates to a large literature on innovation and firm dynamics. Although a few works (including Acs and Audretsch 1987, 1988, Kortum and Lerner 2000, Baumol 2009, Akcigit and Kerr, 2010, and Acemoglu et al., 2013) emphasize heterogeneity in innovation behavior and strategy across firms, we are not aware of other papers in this literature that focus on creative innovations or link this to the incentives and constraints imposed on the behavior of innovators and managers within companies. Nevertheless, within this literature, several papers also emphasize the importance of patent quality. In particular, Trajtenberg (1990), Harhoff et al (1999), Shane and Klock (1997), Sampat and Ziedonis (2004), and Abrams et al. (2013) document a positive relationship between citations and various measures of private or social value.

Finally, the main ideas here have a resonance with the innovation literature investigating disruptive innovations, which follows Christensen's seminal *The Innovator's Dilemma* (1997) and attempts to explain why many companies are unable to maintain their innovativeness following

[4]Relatedly, Sarada and Tocoian (2013) investigate the impact of the age of the founders of a company on subsequent performance using Brazilian data, while Azoulay, Manso and Zivin (2011) show the impact of changes in incentives driven by large academic awards and grants on creativity, and Azoulay, Zivin and Wang (2010) investigate the impact of the death of a very productive co-author on academic productivity.

success. Henderson (2006) discusses the organizational aspects of the innovator's dilemma. Adner and Zemsky (2005) investigate the relationship between disruptive innovations and competition, while King and Tucci (2002) discuss the role of managerial strategies and experience in dealing with these issues. Our potential answer to the innovator's dilemma, consistent both with Arrow's replacement effect and our interaction results, 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.

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 some basic cross-country correlations corroborating the patterns shown in Figure 1. Section 5 presents our main empirical results, which are based on firm-level data. Section 6 concludes.

2 Motivating Theory

In this section, we provide a stylized 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 the profit maximization of the final good sector, 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

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

Equilibrium profits for a product line with technology q_j are

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

where the second line defines π.

2.2 Managers

In addition to workers, the economy is also populated by managers. 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 birth date b. 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 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 Innovation Dynamics

The productivity of each intermediate product is determined by its location along a quality ladder in a given product line. In addition, each leading-edge technology gives the firm an opportunity for further innovation. Two types of innovations are possible:

1. *incremental innovations,* which improve the productivity of a product line within the current *technology cluster.*[5] 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 declines in the number of prior incremental innovations within a technology cluster. In addition, again because these take place within a given technology cluster, they build on a narrow technology base and create improvements 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).

2. *radical innovations*, which combine the current technology of the product line, 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). This combination of knowledge creates a new technology cluster (thus akin to Weitzman's (1998) recombination approach) upon which new incremental innovations can be built. 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.

Firms 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 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, with weight $1-\kappa$, and on average technology, $\bar{q}_t$, with weight κ. When this will cause no confusion, we will suppress the arguments of $\eta_n(q_j, \bar{q}_t)$. 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 be written as

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

[5] Our modeling of technology clusters follows Akcigit and Kerr (2010) and Abrams et al (2013).

If, instead, there is a radical innovation in this particular product line, the innovator will initiate a new technology cluster in a different product line (and will still keep its original product line). First, this ensures a larger improvement on current technology. Second, it generates the ability to start a new series of incremental innovations. However, radical innovations are not directed, and since 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.[6] Thus radical innovations are associated with "Schumpeterian creative destruction." We describe the technology for radical innovations below.

For each of their active product lines, firms hire managers who influence their revenues in two distinct 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).[7] 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.

We assume that there are two types of firms, denoted by $\theta \in \{\theta_H, \theta_L\}$ where $\theta_H > \theta_L = 0$, which means that firms are distinguished by their "corporate culture" determining their openness to disruption and radical innovation. We assume that the arrival rate of a radical innovation for a firm of type θ with a manager with knowledge base $\bar{q}_b$ when the current technology in the economy is $\bar{q}_t$ is given by

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

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 the subscript, here indicates that this is a ratio of two averages). This specification implies that low-type firms, with $\theta_L = 0$, cannot engage in radical innovations.

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

[6] 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.

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

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.[8]

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

$$q_j^0 = (1 + \eta_0)\, q_j,$$

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

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:

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

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

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

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

$\eta_2 < \eta_1$. 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, initiating a new technology cluster.

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 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 also receive more "general" citations. 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.

2.4 Entry, Exit and Firm Dynamics

New firms enter at the exogenous flow rate x, and innovate (improve) over an existing product line uniformly at random. Thus, new firms, upon successful entry, initiate a new technology cluster. Subsequently, a firm's type is determined and remains fixed. We assume that with probability $\zeta \in (0,1)$, each becomes a high type, $\theta_H = \theta$, and with the remaining probability, a low type, $\theta_L = 0$.

As noted above, 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.,

$$\overrightarrow{q_f} \equiv \left\{q_{f,j_1}, q_{f,j_2}, ..., q_{f,j_{m_f}}\right\}, \text{ and } \overrightarrow{n_f} \equiv \left\{n_{f,j_1}, n_{f,j_2}, ..., n_{f,j_{m_f}}\right\},$$

and m_f denotes the number of product lines that firm f is operating.[9] The value function for a low-type firm can be written as

$$rW_L(\overrightarrow{q_f}, \overrightarrow{n_f}) - \dot{W}_L(\overrightarrow{q_f}, \overrightarrow{n_f}) = \sum_{m=1}^{m_f} \left[\begin{array}{c} \max_{a \geq 0} \left\{\pi q_{f,j_m} + \bar{q}_t f(a) - w_{a,t}\right\} \\ +\xi \left[W_L \left(\begin{array}{c} \overrightarrow{q_f} \backslash \{q_{f,j_m}\} \cup \left\{q_{f,j_m} + \eta_{n_{f,j_m}+1}\right\}, \\ \overrightarrow{n_f} \backslash \{n_{f,j_m}\} \cup \{n_{f,j_m} + 1\} \end{array} \right) - W_L(\overrightarrow{q_f}, \overrightarrow{n_f}) \right] \\ -\tau \left[W_L(\overrightarrow{q_f} \backslash \{q_{f,j_m}\}, \overrightarrow{n_f} \backslash \{n_{f,j_m}\}) - W_L(\overrightarrow{q_f}, \overrightarrow{n_f}) \right] \end{array} \right]. \quad (5)$$

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, 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, which we write as 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 $W_L(\overrightarrow{q_f}, \overrightarrow{n_f})$). Finally, 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 below), 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 $\overrightarrow{n_f} \backslash \{n_{f,j_m}\}$).

Note also that in writing this value function, we have simplified the notation with a slight abuse. First, even though the value function depends on calendar time because of its dependence on average technology in the economy, $\bar{q}_t$, we have suppressed time as an argument, and second, we wrote $a \geq 0$ instead of $a \in \mathbb{R}_+ \cup \{\emptyset\}$ to designate the possibility that the firm may end up not hiring a manager.

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

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

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

The intuition for this value function is identical to (5) except for the possibility of a radical innovation. In particular, for each product line m, this high-type firm 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 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

$$
rV_L\left(q_j,n\right)-\dot{V}_L\left(q_j,n\right)=\max_{a\geq 0}\left\{\pi q_j+\bar{q}_t f(a)-w_{a,t}\right\}+\xi\left[V_L\left(q_j+\eta_{n+1},n+1\right)-V_L\left(q_j,n\right)\right]-\tau V_L\left(q_j,n\right), \tag{7}
$$

and

$$
\begin{aligned}
& rV_H\left(q_j,n\right)-\dot{V}_H\left(q_j,n\right) \\
= & \max\left\{\begin{array}{c}\pi q_j+\max_{a\geq 0}\left\{\bar{q}_t f(a)-w_{a,t}+\xi\left[\begin{array}{c}V_H\left(q_j+\eta_{n+1},n+1\right) \\ -V_H\left(q_j,n\right)\end{array}\right]\right\}; \\ \pi q_j+\max_{a\geq 0}\left\{\bar{q}_t f(a)+\Lambda\bar{q}^a\theta_H\mathbb{E}V_H\left(\bar{q}_t\right)-w_{a,t}\right\}\end{array}\right\}-\tau V_H\left(q_j,n\right),
\end{aligned} \tag{8}
$$

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 thus 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.

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:

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

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

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

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

Proposition 2 *The value function of a product line operated by a low-type firm,* (7) *takes the following form*

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

where

$$B \equiv \frac{\xi \eta \alpha A}{r - g + \tau + \xi(1-\alpha)} \text{ and } A \equiv \frac{\pi}{\tau + r}.$$

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 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:

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

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.[10]

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

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

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*

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

[10] 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

$$\tilde{A} = \frac{\pi}{r+\tau},$$

and $\tilde{B}(n)$ *is given by*

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

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*

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

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); intuitively, it is 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:

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

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

Our main results in this subsection fully characterize the stationary equilibrium.

Proposition 4 *At time t managers with $a \leq a^*$ ("young" managers or those with $b \geq b_t^*$) will be hired on product lines for which firms are pursuing radical innovations, which are those operated by high-type firms and that have had more than n^* prior incremental innovations, where n^* is given by (14). Managers with $a > a^*$ ("old" managers or those with $b < b_t^*$) will be hired by firms that undertake incremental innovations. Managerial wages at time t are given by (15).*

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. ■

The implications of changes in Λ are particularly 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 society discouraging radical innovations.

Our empirical work is partly inspired by Proposition 4 (though it does not directly test its results). 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 relationship between society-wide measures of permissiveness to radical innovations (corresponding to Λ) and our measures of creative innovations in the cross-country data. We then turn to the relationship in firm- and patent-level data between manager (CEO) age and creative innovations. In both cases, as our model highlights, the relationship between manager age and creative innovations does not capture a causal effect of manager age on creativity or innovation. In our model, this can be seen from the sorting of young managers to high-type firms, which are the only ones with the ability to undertake radical innovations. But at the same time, for a high-type firm, a young manager contributes to radical innovations, so that there *is* a causal effect of manager age, but this will be confounded with the aforementioned sorting when we look at the correlation between manager age and measures of radical innovation. For this reason, we do not interpret the correlations we present below as causal effects, though at the end of our firm-level analysis we present a tentative strategy for decomposing these correlations between causal and sorting effects.

2.6 General Equilibrium and the Stationary Distribution of Products

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

The aggregate creative destruction rate in the economy results from entry and radical innovations and can be written as

$$\tau = x + M \int_0^{a^*} \Lambda \bar{q}^a \theta dF(a),$$

where x is the entry rate, $F(a)$ denotes the stationary distribution of manager age, and a^* is the threshold below which managers are hired by firms to do radical innovation. We can further split this aggregate creative destruction rate into its components coming from low- and high-type firms:

$$\tau^L = x(1-\zeta) \quad \text{and} \quad \tau^H = x\zeta + M \int_0^{a^*} \Lambda \bar{q}^a \theta_H dF(a).$$

Clearly $\tau = \tau^H + \tau^L$. Note that low-type firms generate creative destruction only when they initially enter the economy (since they do not engage in radical innovation).

Let us denote the fraction of product lines occupied by high- and low-type firms with n prior incremental innovations by, respectively, μ_n^H and μ_n^L (these are not functions of time as we are focusing on a stationary equilibrium). Naturally,

$$\sum\nolimits_{n=0}^{\infty} \left[\mu_n^H + \mu_n^L\right] = 1.$$

The invariant step size distribution is determined by the following flow equations for high types

$$\begin{array}{rcllr} \underline{\text{Outflow}} & & \underline{\text{Inflow}} & & \\ (\tau+\xi)\mu_0^H & = & \tau^H & \text{for} & n=0 \\ (\tau+\xi)\mu_n^H & = & \xi\mu_{n-1}^H & \text{for} & n^* > n > 0 \\ \tau\mu_{n^*}^H & = & \xi\mu_{n^*-1}^H & \text{for} & n = n^* \\ \mu_n^H & = & 0 & \text{for} & n > n^* \end{array}.$$

Intuitively, entry into the state of high-tech product lines with $n=0$ is driven by radical innovation from high-type firms, which takes place at the flow rate τ^H. Exit from this state takes place when the current firm engages in an incremental innovation, at the flow rate ξ or when there is creative destruction (from both high- or low-type firms), which takes place at the aggregate creative destruction rate, τ. Entry and exit into other states have similar intuitions, except that there is no entry into states with $n > n^*$ since high-type firms switch to radical innovation at $n = n^*$.

The flow equations for the low-type product lines can be written similarly as:

$$\begin{array}{rcllr} \underline{\text{Outflow}} & & \underline{\text{Inflow}} & & \\ (\tau+\xi)\mu_0^L & = & \tau^L & \text{for} & n=0 \\ (\tau+\xi)\mu_n^L & = & \xi\mu_{n-1}^L & \text{for} & n>0 \end{array}.$$

These can be solved for the following geometric distributions for high- and low-type firms:

$$\mu_n^L = \left[\frac{\xi}{\tau+\xi}\right]^n \frac{\tau^L}{\tau+\xi} \text{ for } n \in \mathbb{Z}_+, \quad \text{and} \quad \mu_n^H = \begin{cases} \left[\frac{\xi}{\tau+\xi}\right]^n \frac{\tau^H}{\tau+\xi} \text{ for } n < n^* \\ \left[\frac{\xi}{\tau+\xi}\right]^n \frac{\tau^H}{\tau} \text{ for } n = n^* \end{cases}.$$

To derive the aggregate growth rate, let us also combine (1) with (2), which gives

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

The growth rate of the economy is then equal to the growth of the average quality $\bar{q}_t$. After a small time interval $\Delta t > 0$, the average quality evolves according to the following law of motion:

$$\bar{q}_{t+\Delta t} = \bar{q}_t + \eta\bar{q}_t\left[x + \mu^H_{n^*}Q\Lambda\theta\right]\Delta t + \bar{q}_t\xi\eta\Delta t\left[\sum_0^{n^*}\mu^H_n\alpha^n + \sum_0^{\infty}\mu^L_n\alpha^n\right] + o(\Delta t),$$

where $Q \equiv \frac{1}{F(a^*)}\int_0^{a^*}\bar{q}^a dF(a)$ is the average productivity gap of the managers that are hired for radical innovations, and $o(\Delta t)$ denotes terms that are second order in Δt. Then the growth rate in the stationary equilibrium can be obtained as:

$$g = \eta\left[x + \mu^H_{n^*}Q\Lambda\theta\right] + \xi\eta\left[\sum_0^{n^*}\mu^H_n\alpha^n + \sum_0^{\infty}\mu^L_n\alpha^n\right].$$

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:

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

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 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:

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

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 $\kappa < 1$), 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.

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.[11]

Compustat North American Fundamentals The Compustat data for publicly traded firms in North America are from Wharton Research Data Services. This dataset contains a detailed list of balance sheet items reported by the companies annually between 1974 and 2006. It contains 29,378 different companies, and 390,467 *company* × *year* observations. The main variables of interest are

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

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.[12]

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.

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).[13] 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.

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

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

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, 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.

Other Data Sources The average years of schooling in secondary education is used as a proxy of the human capital of a country, retrieved from the Barro-Lee dataset (Barro, Lee forthcoming)[14]. Real GDP per capita numbers and R&D intensity come from the World Bank's World Development Indicators database.

In our baseline analysis, we focus on citation and patents between 1995 and 2000 (with patents classified according to their year of *application*). This is motivated by our wish to construct a balanced panel of firms in our baseline firm-level analysis, where we use a single observation per firm (extending the beginning of our sample to the earliest date at which we have manager age, 1992, significantly reduces our sample). We also stop the sample in 2000 (or 2002), so that we have a sufficient subsequent window during which to measure citations. We show below that our results are robust to extending the sample.

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 *country* $\times$ *year* and *company* $\times$ *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.

[14] http://www.barrolee.com/data/dataexp.htm

All inventors are ranked according to this score, and the top 5% are considered to be superstar inventors. The superstar fraction of a country or company in a given year is calculated as the fraction of patents with superstar inventors in that year (if a patent has more than one inventors, it gets a fractional superstar designation equal to the ratio of superstar inventors to the total number of inventors on 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 countries, companies or innovators have the tendency for generating "tail innovations" with very high citations.

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 originality and generality 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 cross-country, balanced firm and unbalanced firm 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 56.1 in our cross-country sample and 52.3 in our firm-level (balanced or unbalanced Compustat) 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.

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

4 Cross-Country Correlations

We start our empirical analysis by providing a few more details on the cross-country patterns shown in Figure 1. Though our main evidence comes from firm-level and patent-level regressions presented in the next section, the results in this section show that the cross-country patterns discussed in the Introduction are also fairly robust.

We should note at this point that, as already mentioned, the interpretation of cross-country and firm-level results could in fact be different. At the firm level, as our theory highlighted and we will emphasize again below, the age of managers is in part an indicator for a certain type of firm or "corporate culture" (corresponding to our high-type firms) that can undertake radical innovations. Nevertheless, conditional on having such a firm, a young manager does also contribute to radical innovations (because of his more recent knowledge stock). At the country level, however, manager age, like our other measures, is likely to have much of its impact on the creativity of innovations entirely through institutions, attitudes and values of the society within which it is correlated.

Table 2 reports OLS results from cross-country regressions of the following form:

$$y_c = \alpha I_c + \mathbf{X}_c' \boldsymbol{\beta} + \varepsilon_c, \tag{16}$$

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 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 error term. 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 2 include one observation per country, report robust (against heteroscedasticity) standard errors, and use the total number of patents as weights. The weighted specification is 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. Appendix Table A1 shows the distribution of total number of patents across countries.[15]

Panel A of Table 2 focuses on Hofstede's individualism index and contains results for our four main measures of creative innovations. Column 1, for example, has an estimate of 4.97 (standard error = 2.46) in the first row for innovation quality (average number of citations). This estimate thus shows a positive association between the individualism index and the average number of citations per patent. The other rows show the effect of log GDP per capita, average years of secondary schooling and log total number of patents. GDP per capita and average years of secondary schooling are not significant, a pattern common with most other specifications we report in this table, while log patent count is significant and indicates that countries that have more patents also tend to have more citations per patent. The quantitative magnitude of the correlation between individualism and innovation quality is sizable. 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 the weighted sample mean (14.5). In fact, the R^2 at the bottom of the panel indicates that, despite its parsimony, this specification explains 73% of the cross-country variation in average citations per patent.

The other columns of the table show the correlation between the individualism index and our other measures, in particular the superstar fraction of innovations, tail innovations, and generality. In each case, there is a fairly strong and highly significant correlation between individualism and these indices (the estimate of the coefficient on individualism is significant at less than 1% in all of these cases). The magnitudes for the superstar fraction and tail innovation variables are significantly larger than in the first column: in the former case, moving from the country at the

[15] 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.

25th percentile of individualism in our sample to that at the 75th percentile increases the superstar fraction by 80% relative to the weighted sample mean (6.68), whereas in the latter case, it increases by 67% relative to the weighted sample mean (1.92).

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

In Panel C, we turn to our measure of manager age. This reduces our sample from 50 to 37 due to the more limited availability of this measure. As we argued above, we believe this is a good proxy for openness to disruption because only companies and countries that are open to new and potentially disruptive ideas, innovations and practices enable young managers to rise up to the highest positions rather than relying on age, experience and slow movements within the hierarchy. The patterns are very similar in this case also, with a strong correlation between average manager age and all four of our measures of creative innovations. The quantitative magnitudes are also broadly similar in this case. 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). More relevant for comparison with our firm-level and patent-level results might be to look at how much a one-year change in manager age impacts our various measures of creative innovations. Here the results are again reasonable. For example, such a change would increase 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).[16]

Tables 3 and 4 probe the robustness of the cross-country relationships reported in Figure 1 and Table 2. Table 3 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 4, 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

[16] 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.

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 4 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.[17]

5 Firm-Level and Patent-Level Results

Our main empirical results exploit firm-level variation in manager age across Compustat companies. Though several other factors determine manager age, we believe that, as with the cross-country variation, a younger manager/CEO reflects a greater openness to disruption. Motivated by this reasoning, in this section we show the relationship between firm-level measures of creative innovations and manager age.

Two caveats are important at this point. First, 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). Second, consistent with our theoretical framework where different types of firms select into hiring of young managers (recall Propositions 2 and 5), we do not interpret the results we report in this section as the "causal effect" of manager age on creative innovations. Instead, we believe that manager age at least partly proxies for the company's overall openness to disruption. At the end of this section, we turn to a more direct investigation of the effect of manager age on creative innovations.

5.1 Main Results

Our main results are presented in Table 5. Our estimating equation is similar to (16),

$$y_f = \alpha m_f + \mathbf{X}'_f \boldsymbol{\beta} + \delta_{i(f)} + \varepsilon_f, \tag{17}$$

where y_f is one of our measures of creative innovations (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

[17] 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 motivation rather than as our main empirical focus.

time window (we do not have measures of the human capital of the firm's employees).[18] Controlling for firm age is particularly important, since we would like 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 included in all regressions so that the comparisons are always across firms within a fairly narrow industry. 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). We first exploit only cross-sectional information, so our regressions have one observation per firm and are weighted with the total patent count of the firm. All standard errors are again robust. Different columns of Table 5 correspond to our four different measures of creative innovations, now constructed at the firm level.

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 sizable and comparable to the quantitative magnitudes we obtained in the cross-country data when using average manager age. For example, the effect of a one-year increase in CEO age is to raise average citations by 0.278, which is about 60% of the magnitude of the cross-country relationship between average manager age and innovation quality.[19]

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 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 10 below). Columns 3 and 4 show even more precisely estimated (significant at 1% or less) and economically large relationships with our measures of tail

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

[19] If, instead, we use the average age of top management, the quantitative impact rises to 0.418; which is more comparable to the cross-country magnitude. The quantitative implications of moving from the 75th percentile of the manager age distribution to the 25th percentile are also more similar to the cross-country magnitudes.

innovations and generality.

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. Though this relationship may not be causal (or may reflect the impact of CEO age working through other channels than openness to disruption), it is both new and quite consistent with our theoretical expectations. We will next see that it is also quite robust.

5.2 Robustness

Tables 6 and 7 probe the robustness of our firm-level results in different dimensions. Table 6 looks at the same alternative measures of creative innovations we studied in Table 3 in the cross-country context (recall that 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 5, except that the relationship is no longer statistically significant with the alternative measure of the superstar fraction.

Table 7 looks at several different robustness exercises. Panel A replaces the four-digit SIC dummies with three-digit dummies, 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 5, but somewhat 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, includes R&D intensity (R&D to sales ratio) in the previous specification.[20] This is intended, as in the cross-country context, 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 5.

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 5 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

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

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.

5.3 Panel Results

As noted above, our baseline (balanced) sample uses one observation per firm and focuses on 1995-2000. In this section, we investigate several additional issues. First, we show that our results hold if we look at a considerably larger sample spanning a longer time period. Second, and more importantly, we also show that, though naturally much noisier, the results are also consistent when we exploit within-firm variation in the age of the CEO. Third, and consistent with these within-firm results, we provide some evidence that it is the age of the current CEO that seems to matter most for the creativity of innovations. Fourth, we also use these results to shed some preliminary light on the relative importance of the impact of manager age on creative innovations vs. sorting of young managers across different types of firms.

With this objective in mind, in Table 8 we start with our baseline balanced sample, but now we 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. 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 5, and indeed, the coefficient estimates and standard errors are very similar (though they are not identical since the covariates are now time-varying).[21]

Panel B turns to an unbalanced panel and 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 1992-2002 (we cannot go before 1992 because of the lack of data on manager age, and we prefer not to go beyond 2002, as this would make the citation window too short and thus our measures much less reliable). The resulting sample has 6074 observations (or 5268 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 1208 (from 279) and

[21] 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.

the addition of four 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.[22] Hence, this is a very demanding specification investigating whether in years where a firm has a younger CEO, it tends to have more creative innovations, and this motivates our choice of focusing on the 1992-2002 sample for this exercise. In addition to throwing away all of the (potentially useful) 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.[23] 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 generality, have the same sign and are statistically significant as in our baseline results in Table 5. For innovation quality, the magnitude of the estimate is about 18% smaller than the specification without fixed effects in Panel B (e.g., -0.163 vs. -0.200), and for superstar fraction and tail innovations, it is about 40% of the magnitude in Panel B.

One concern is that the current CEO may have only a limited impact on innovations patented in a given year that naturally build on research that had been done many years prior. Counteracting this is that the current CEO may set the strategy that determines which innovations are pursued and marketed and seek patent protection during his or her reign. A natural question is therefore whether it is current CEO age or lagged CEO age that is more important for the creativity of innovations. This is investigated in Panel D. Perhaps again somewhat surprisingly, these results show that it is current CEO age that is the key correlate of our measures of creativity of innovations (except for generality where we could not find any clear pattern). In fact, in all three specifications (innovation quality, superstar fraction and tail innovations), current CEO age is statistically significant with a magnitude close to that in Panel C, while lagged CEO age is not.

A related concern is whether we are partially capturing the persistent effects of past innovations.

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

Observe that in our model a high-type firm will pursue an incremental innovation strategy for a while and then switch to a radical innovation strategy while simultaneously changing its manager to a younger one. In this case, the fixed effect estimator may provide an upper bound on the impact of a younger manager on creative innovations for high-type firms.

[23] 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.

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 very similar to those in Panel C.[24]

It is noteworthy that the inclusion of firm fixed effects reduces the coefficient estimate on CEO age but only slightly (e.g., for innovation quality, from -0.200 in Panel B to -0.163 in Panel C). This suggests that part of the relationship we observe in the cross-section is due to the sorting of younger CEOs to firms that are already more likely to engage in creative innovations, but there is also a significant impact of a younger CEO on creative innovations for the same firm. This is also confirmed by the relative explanatory powers of CEO age and firm fixed effects. If we just include log employment, log sales, log patents, firm age, four-digit SIC dummies and application year dummies (that is, the same variables as in our baseline regressions except CEO age), the R^2 of the regression for innovation quality is 0.49. Once we add CEO age, this increases to 0.64. When we also add firm fixed effects, it further increases to 0.77. This suggests that the explanatory powers of CEO age and firm fixed effects are roughly comparable,[25] even though firm fixed effects here capture not just the sorting channel, so this comparison would provide an upper bound on the quantitative magnitudes of the sorting channel vs. the direct impact of CEO age on creative innovations. The pattern is similar for the other measures of creative innovations. Overall, we tentatively conclude that both the sorting effects and the direct effect of manager age on creativity of innovations are present and sizable.

In the next subsection, we pursue another approach to shed more light on the effect of manager age on the creativity of innovations.

5.4 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

[24] 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.

[25] Of course, the order in which these two sets of variables are added is important. If firm fixed effects are added without CEO age, then the R^2 is 0.72.

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.

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

$$y_{ift} = \phi I_{ift} + \alpha m_{ft} + \mathbf{X}'_{ift}\boldsymbol{\beta} + \delta_f + \gamma_i + d_t + \varepsilon_{ift}. \tag{18}$$

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 are 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 described below. All specifications also control for a full set of year effects, denoted by d_t, and ε_{ift} is the error term.[26]

The results from the estimation of (18) are reported in Table 9. In Panel A we focus on a specification similar to the regressions with firm fixed effects reported in Table 8. 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 1992-2002 and includes the same set of controls as in Table 8 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 8, though a little smaller. In column 1, for instance, we see an estimate of -0.108 (standard error $= 0.040$) compared to -0.163 in Table 8. 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

[26] 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.

marginally significantly in these specifications. Finally, we also show results with the generality index, even though the results in Table 8 already indicated that, with firm fixed effects included, there is no longer a significant relationship between CEO age and the generality index, and this lack of relationship persists for all of the estimates we report in Table 9 (and for this reason, though we do show them for completeness, we will not discuss them in detail).

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;[27] 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.[28] 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.[29] 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 -0.236 (standard error = 0.027), 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 1992-2002. 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 1985-2002. 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

[27] 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).

[28] This corresponds to 374 separate technology classes and is roughly at the same level of disaggregation as the 375 SIC dummies we used in the firm-level analysis in Tables 5-7.

[29] 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.

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); 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 manager/CEO age has an effect on the creativity of innovations that goes beyond the sorting of managers to firms. The quantitative magnitudes are smaller than those in our baseline firm-level regressions, which suggests that some of the firm-level relationship was indeed capturing the endogenous sorting of younger managers toward firms that are more open to disruption. Our next results, reported in Table 10, 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 (18) 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 10 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 9. 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.[30]

5.5 Stock of Knowledge, Opportunity Cost and Creativity of Innovations

Finally, Table 11 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

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

companies with high sales).

This is a demanding, as well as crude, test, since neither proxy is perfect, and moreover, log 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 11, which uses the same unbalanced sample with annual observations as in Table 8 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.[31] 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).

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.

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

[31] As noted above, the main effects are evaluated at the sample mean and are typically close to the estimates reported in Table 5.

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 country, firm and patent level, we find fairly consistent and robust correlations between openness to disruption and creative innovations. We also show that these relationships are generally robust.

Our theoretical model also 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. We reported results from our firm-level data confirming this pattern as well.

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 various other firm-level or cross-country characteristics on the creativity of innovations is a natural direction, which could both rely on instrumental-variables strategies and exploit the structural relationships implied by a more detailed microeconomic model. 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

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

$$\begin{aligned} r\left[Aq_j + B\bar{q}_t\alpha^n\right] - Bg\bar{q}_t\alpha^n &= \pi q_j + \xi\left[\begin{array}{c} Aq_j + A\eta\bar{q}_t\alpha^{n+1} + B\bar{q}_t\alpha^{n+1} \\ -Aq_j - B\bar{q}_t\alpha^n \end{array}\right] \\ &\quad -\tau Aq_j - \tau B\bar{q}_t\alpha^n \end{aligned}$$

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

$$rA = \pi - \tau A$$

and

$$rB - Bg = \alpha\xi A\eta + \xi B(\alpha - 1) - \tau B.$$

Solving these equations for A and B 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

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

which implies

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

Once again equating coefficients, we obtain

$$\tilde{A} = \frac{\pi}{r+\tau}$$

and

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

Now define $\hat{B}(n)$ implicitly as

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

This function can be written as

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

where $\beta = \frac{\xi}{(r-g+\tau+\xi)}$. From standard dynamic programming arguments (e.g., Theorem 4.7 in Stokey and Lucas, 1989), $\hat{B}(n)$ is strictly decreasing.

This implies that there exists n^* such that firms with $n < n^*$ will undertake incremental innovation and will 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. Define $q^*_{n,t}$ as the quality level that makes a cluster that has had n prior innovations so far at time t just indifferent between radical incremental innovation (for a high-type firm). Then we have that

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

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

$$V_H\left(q^*_{n,t} + \eta_{n+1,t}(q^*_{n,t}), n+1\right) - V_H\left(q^*_{n,t}, n\right) = \frac{\Lambda \bar{q}^a \theta}{\xi} \mathbb{E} V_H(t) \tag{19}$$

and

$$(r+\tau) V_H\left(q^*_{n,t}, n\right) - \dot{V}_H\left(q^*_{n,t}, n\right) = \pi q^*_{n,t} + \Lambda \bar{q}^a \theta \mathbb{E} V_H(t) \tag{20}$$

Now we will consider two alternative cases:

Case 1:

$$q^*_{n+1,t} \geq q^*_{n,t} + \eta_{n+1,t}(q^*_{n,t}). \tag{21}$$

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 (by mistake or off-the-equilibrium path), then subsequently it will still want to immediately switch to radical innovation.

Under this case, we have

$$(r+\tau) V_H\left(q^*_{n,t} + \eta_{n+1,t}(q^*_{n,t}), n+1\right) - \dot{V}_H\left(q^*_{n,t} + \eta_{n+1,t}(q^*_{n,t}), n+1\right) = \pi q^*_{n,t} + \pi \eta_{n+1,t}(q^*_{n,t}) + \Lambda \bar{q}^a \theta \mathbb{E} V_H(t). \tag{22}$$

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 (19) with respect to time, we have

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

where we have used the fact that in a stationary equilibrium $\mathbb{E} V_H(t)$ grows at the rate g. We can now subtract (20) from (22) to obtain:

$$(r+\tau)[V_H\left(q^*_{n,t} + \eta_{n+1,t}(q^*_{n,t}), n+1\right) - V_H\left(q^*_{n,t}, n\right)] = \pi \eta_{n+1,t}(q^*_{n,t}) + \frac{\Lambda \bar{q}^a \theta}{\xi} g \mathbb{E} V_H(t) \tag{23}$$

Then, combining (19) and (23) we can derive

$$\eta_{n+1,t}(q^*_{n,t}) = \frac{r-g+\tau}{\pi \xi} \Lambda \bar{q}^a \theta \mathbb{E} V_H(t). \tag{24}$$

In this case, for all q less than $q^*_{n,t}$, it is optimal to switch to radical innovation. Moreover, $q^*_{n,t}$ is increasing in n. We next derive the condition under which (21) indeed applies.

Defining $v_t \equiv \frac{r-g+\tau}{\pi \xi} \Lambda \bar{q}^a_a \theta \mathbb{E} V_H(t)$, equation (24) implies:

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

or

$$q^*_{n,t} = \frac{v(t)/\eta\alpha^{n+1} - \kappa\bar{q}_t}{1-\kappa}, \tag{25}$$

and similarly

$$q^*_{n+1,t} = \frac{v(t)/\eta\alpha^{n+2} - \kappa\bar{q}_t}{1-\kappa}. \tag{26}$$

Combining these equations, we obtain that (21) is satisfied if

$$(1-\kappa)\eta\alpha^{n+2} + \alpha \leq 1. \tag{27}$$

Case 2:

$$q^*_{n+1,t} - \eta_{n+1,t}(q^*_{n,t}) < q^*_{n,t}. \tag{28}$$

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 (21).

In this case, start with $q^*_{n+1,t}$, which satisfies (22). Under condition (28), $q^*_{n,t}$ satisfies (20), so we again arrive at (19) and (24). This gives the same expression as for $q^*_{n,t}$ and $q^*_{n+1,t}$ as in (25) and (26). Thus the condition for (28) to be satisfied, with an identical argument, is

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

which is the complement of (27), and thus establishes that (25) still applies, and thus for all q less than $q^*_{n,t}$, it is optimal to switch to radical innovation, and $q^*_{n,t}$ is increasing in n. 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 WP 19763.

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

Acs, Z. J. and D. B. Audretsch (1987). Innovation, Market Structure, and Firm Size. *Review of Economics and Statistics*, 567-574.

Acs, Z. J. and D. B. Audretsch (1988). Innovation in Large and Small Firms: An Empirical Analysis. *American Economic Review*, 678-690.

Adner, R. and P. Zemsky (2005). Disruptive Technologies and the Emergence of Competition. *RAND Journal of Economics*, 229-254.

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 WP 16443.

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.

Baumol, W. J. (2009). Small Enterprises, Large Firms and Growth. In *The Role of SMEs and Entrepreneurship in a Globalised Economy*, ed. by A. Lundström, Ministry of Education and Research Sweden, 59 – 72.

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

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

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

Cyert, R. M. and G. March, James (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 WP 18470.

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

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 WP 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 WP 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.

Henderson, R. (2006). The Innovator's Dilemma as a Problem of Organizational Competence. *Journal of Product Innovation Management,* 23 (1), 5-11.

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.

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.

King, A. A. and C. L. Tucci (2002). Incumbent Entry into New Market Niches: The Role of Experience and Managerial Choice in the Creation of Dynamic Capabilities. *Management Science,* 48 (2), 171-186.

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.

Kortum, S. and J. Lerner (2000). Assessing the Contribution of Venture Capital to Innovation. *RAND Journal of Economics*, 674-692.

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.

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, H. and M. Klock (1997). The Relation between Patent Citations and Tobin's Q in the Semiconductor Industry. *Review of Quantitative Finance and Accounting,* 9 (2), 131-146.

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.

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.

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

Weitzman, M. L. (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
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
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, 1992-2002)			
CEO age	6074	55.5	6.83
average manager age	6074	52.3	4.43
innovation quality	6074	17.3	10.6
superstar fraction	6074	10.5	11.0
tail innovation	5268	2.79	3.78
generality	5697	19.9	9.23
log patents	6074	5.66	1.60
log employment	6074	3.73	1.50
log sale	6074	4.15	1.60
firm age	6074	34.9	16.0

- Table 1 continued on next page -

Panel B: 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

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 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

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 Cross-Country Regressions

	Innovation Quality	Superstar Fraction	Tail Innovation	Generality
	Panel A: 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 B: Uncertainty Avoidance			
uncertainty aversion	-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
	Panel C: 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

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 3: Cross-Country Regressions (Alternative Measures)

	Innovation Quality (5 years)	Superstar Fraction (Best Patent)	Tail Innovation (90/50)	Originality
	Panel A: 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 B: Uncertainty Avoidance			
uncertainty aversion	-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
	Panel C: 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

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

	Innovation Quality	Superstar Fraction	Tail Innovation	Generality
	Panel A: 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 B: Uncertainty Avoidance			
uncertainty aversion	-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
	Panel C: 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

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 5: 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 6: 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 7: 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 5 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 8: 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, 1992-2002				
average CEO age	-0.200	-0.326	-0.129	-0.162
	(0.077)	(0.130)	(0.042)	(0.047)
R^2	0.63	0.55	0.30	0.74
N	6,074	6,074	5,268	5,697
Panel C: CEO Age (Fixed Effects), Unbalanced Firm Sample, 1992-2002				
CEO age	-0.163	-0.130	-0.044	0.031
	(0.050)	(0.054)	(0.021)	(0.033)
R^2	0.77	0.81	0.52	0.84
N	6,074	6,074	5,268	5,697
Panel D: CEO Age and Lagged CEO Age (Fixed Effects), Unbalanced Firm Sample, 1993-2002				
CEO age	-0.137	-0.104	-0.035	0.028
	(0.042)	(0.041)	(0.021)	(0.030)
lagged CEO age	-0.091	-0.074	-0.020	0.017
	(0.068)	(0.064)	(0.030)	(0.040)
R^2	0.79	0.82	0.54	0.84
N	4,858	4,858	4,292	4,562
Panel E: CEO Age and Lagged Dependent Var (Fixed Effects), Unbalanced Firm Sample, 1993-2002				
CEO age	-0.103	-0.079	-0.035	0.035
	(0.034)	(0.037)	(0.017)	(0.029)
lagged dependent variable	0.417	0.407	0.162	0.147
	(0.037)	(0.047)	(0.047)	(0.038)
R^2	0.84	0.86	0.55	0.84
N	5,522	5,522	4,510	5,091

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 9: Patent-Level Panel Regressions

	Innovation Quality	Tail Innovation (Above 99)	Tail Innovation (Above 90)	Generality
Panel A: CEO Age, Unbalanced Firm Sample, 1992-2002				
CEO age	-0.108	-0.293	-1.066	0.029
	(0.040)	(0.132)	(0.422)	(0.026)
R^2	0.11	0.03	0.07	0.11
N	311,216	311,216	311,216	261,266
Panel B: Inventor Age, Unbalanced Firm Sample, 1992-2002				
inventor age	-0.236	-0.444	-2.924	-0.019
	(0.027)	(0.120)	(0.327)	(0.022)
R^2	0.14	0.03	0.09	0.15
N	311,216	311,216	311,216	261,266
Panel C: Inventor Age, Extended Sample, 1985-2002				
inventor age	-0.228	-0.376	-2.872	-0.017
	(0.022)	(0.075)	(0.296)	(0.017)
R^2	0.16	0.05	0.10	0.14
N	562,552	562,552	562,552	462,313
Panel D: Inventor Age, Extended Sample, 1985-2002				
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, 1992-2002				
inventor age	-0.235	-0.443	-2.919	-0.019
	(0.027)	(0.120)	(0.327)	(0.022)
CEO age	-0.111	-0.302	-1.071	0.029
	(0.038)	(0.126)	(0.402)	(0.023)
R^2	0.14	0.03	0.09	0.15
N	311,216	311,216	311,216	261,266

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). 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 10: Inventor Age and CEO Age,
Unbalanced Firm Sample, 1992-2002

	Inventor age	Inventor age
	(1)	(2)
CEO age	0.012	0.012
	(0.006)	(0.006)
R^2	0.11	0.13
N	311,216	311,216

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 dummies, 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 11: Stock of Knowledge, Opportunity Cost, and Creative Innovations,
Unbalanced Firm Sample, 1992-2002

	Innovation Quality	Superstar Fraction	Tail Innovation	Generality
CEO age	-0.178	-0.207	-0.073	-0.049
	(0.030)	(0.031)	(0.014)	(0.018)
log sales	1.477	2.188	0.224	1.309
	(0.485)	(0.672)	(0.219)	(0.352)
log patent	-0.439	-0.268	0.240	-0.051
	(0.208)	(0.284)	(0.107)	(0.158)
average CEO age $\times$ log patent	-0.003	-0.067	-0.028	-0.036
	(0.015)	(0.023)	(0.009)	(0.012)
average CEO age $\times$ log sales	0.036	0.086	0.032	0.045
	(0.019)	(0.025)	(0.010)	(0.012)
R^2	0.64	0.55	0.31	0.74
N	6,074	6,074	5,268	5,697

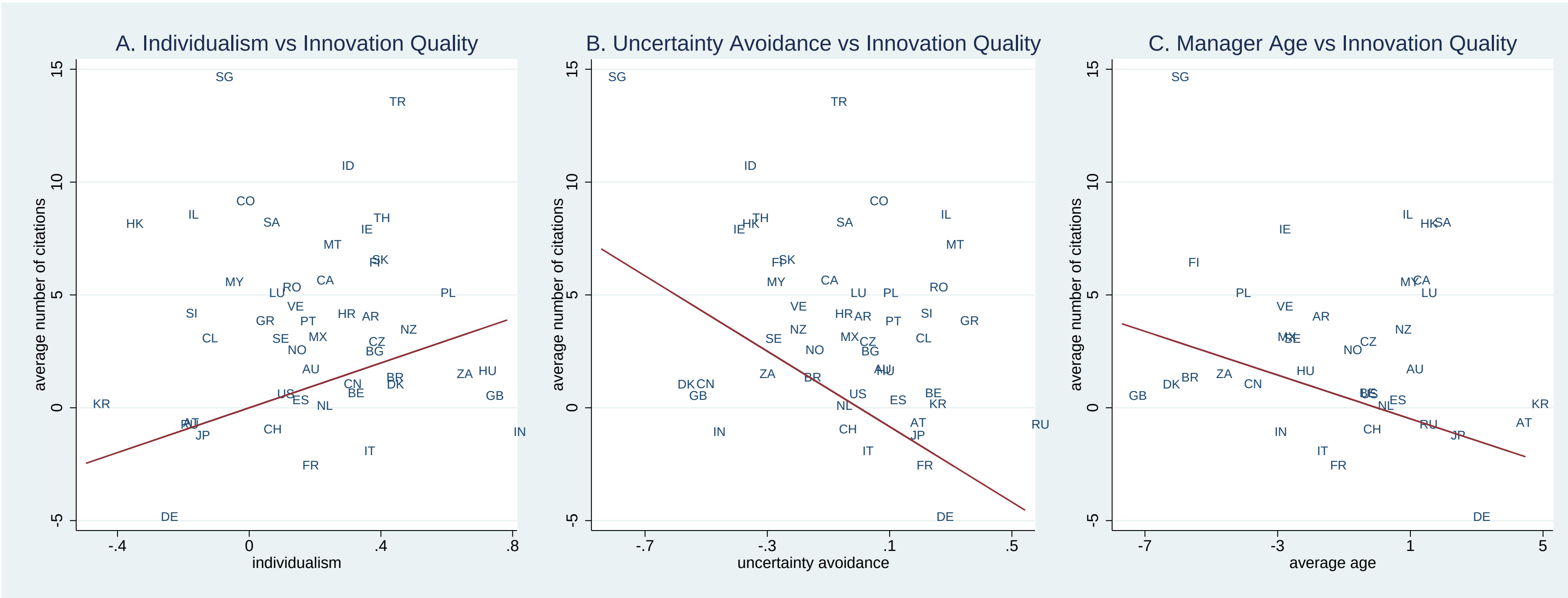
Notes: Weighted firm-level panel regressions with annual observations for the unbalanced firm panel, 1992-2002, with number of patents (in that year) as weights. The dependent variables are innovation quality, superstar fraction, tail innovation, and generality (the last three are multiplied by 100 to 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 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: Openness to Disruption Measures vs Innovation Quality



Notes: Residual plots from a weighted regression of innovation quality (average number of citations per patents) on Hofstede's individualism index, Hofstede's uncertainty avoidance index, and our average manager age variable on log income per capita, averages of secondary years of schooling, and log total number of patents, with total number of patents as weights for 1995-2000. See text and notes to Table 1 for variable definitions.