Quasi-Experimental Estimates of Rent Sharing: Evidence from Administrative Tax Data*

Patrick Kline	Neviana Petkova	Heidi Williams	Owen Zidar
UC-Berkeley	US Department of Treasury	MIT	Chicago Booth

March 25, 2016

Abstract

We develop a new firm-level linkage between patent applications submitted to the US Patent and Trademark Office (USPTO) and US Treasury tax files. Using these data, we investigate how patent-generated monopoly rents impact US firms and workers. We first document that firm profits rise dramatically after a patent application is filed, while worker compensation rises in advance of the application but stabilizes once the application is submitted. Both of these post-application trends in firm and worker outcomes are more pronounced among applications which are ex ante more valuable, suggesting a causal relationship between the underlying innovation which gives rise to the patent application and these changes in firm and worker outcomes. We then explore the impacts of patent grants – as opposed to the innovation underlying the patent application – by instrumenting for patent grants with the leniency of the assigned patent examiner. While our preliminary instrumental variable results are imprecise, they currently suggest that patent grants per se have mild positive effects on firm profitability and worker compensation.

Very preliminary: Please do not cite nor circulate

*We are very grateful to Alex Fahey, Sam Grondahl, and John Wieselthier for excellent research assistance. Research reported in this publication was supported by the National Institute on Aging and the NIH Common Fund, Office of the NIH Director, through Grant U01-AG046708 to the National Bureau of Economic Research (NBER); the content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or NBER. This work/research was also funded by the Ewing Marion Kauffman Foundation; the contents of this publication are solely the responsibility of the Grantee. Financial support from NSF Grant Number 1151497, the Kathryn and Grant Swick Faculty Research Fund, and the Booth School of Business at the University of Chicago are also gratefully acknowledged. Opinions expressed in the paper do not represent the views and policies of the US Department of the Treasury.

Contact: pkline@berkeley.edu, Neviana.Petkova@treasury.gov, heidiw@mit.edu, Owen.Zidar@chicagobooth.edu.

A classic question in economics is the extent to which firms share rents with their workers. In the textbook model of a competitive labor market, a worker's wage depends only on her productivity – the profitability of her employer is irrelevant to the wage setting process. Deviations from the competitive model of wage setting have important implications for a host of topics in labor economics, public finance, and macroeconomics including the effects of productivity dispersion on wage inequality (Barth, Bryson, Davis and Freeman, 2011), the incidence and employment effects of the minimum wage (Flinn, 2006, 2010), optimal corporate and income taxation (Piketty, Saez and Stantcheva, 2014), and the propagation and incidence of macroeconomic shocks in a frictional environment (Shimer, 2005; Hagedorn and Manovski, 2008).

In this paper, we investigate how the temporary rents generated by patent grants affect the profitability of firms and the wages of workers. Our analysis builds on the seminal contribution of Van Reenen (1996) who first studied the impact of technological innovations on worker wages in a panel of UK firms. He estimates an elasticity of worker wages with respect to quasi-rents of 0.29 – a relatively large estimate in comparison to other studies in the literature (see Card et al. (2016) for a review). One concern with his analysis is that technological innovations tend to reflect substantial and purposeful research investments made by firms (see, e.g., the discussion in Hall, Jaffe and Trajtenberg (2005)). Hence, observational variation in the timing of innovations across firms may be driven in part by confounding trends, and many aspects of the firm beyond simply rents – for example, the firm's production process – may change around the time of the innovation. A second limitation is that Van Reenen (1996) only had access to data on firms' aggregate outcomes, which raises the possibility that some of the wage increases he estimates could be driven by skill upgrading behavior by firms.

In order to investigate this question, we first develop a new firm-level linkage between published US patent applications (both successful and unsuccessful) and US Treasury tax files. Specifically, we leverage US Treasury tax files providing longitudinal data on firms' Forms 1120 (C-Corporations), 1120-S (S-Corporations), and 1065 (Partnerships), linked to workers' W2 Forms, over the years 1997-2014. This new linkage enables us to – for the first time – document descriptive facts on the landscape of patenting activity by US firms, not limited (as in the past literature, e.g. Hall, Jaffe and Trajtenberg (2005)) to measuring patenting activity among large publicly traded firms that cover a selected and non-representative sample of US workers. Broadly stated, our empirical approach first estimates how patents affect firm profitability, and then estimates how patents affect the compensation of current and future employees. By scaling the compensation estimate by the firm profitability estimate, we can obtain an estimate of the elasticity between wages and quasi-rents.

Our empirical analysis offers two main contributions relative to the prior work which has empirically tested for evidence of rent sharing. First, we use patent applications as well as patent grants as measures of technological innovations. We use an event study framework to investigate how firm and worker outcomes change around the time a patent application is filed. We argue that these estimates are likely to capture both changes induced by

the underlying innovation which gives rise to the patent application, and some effect of the patent application filing. In order to investigate patent grants – as opposed to the innovation underlying the patent application – as a quasi-experimental source of variation of firm rents, we use the "leniency" of the patent examiner assigned to review the patent application at the US Patent and Trademark Office (USPTO) as a source of exogenous variation in the timing of patent grants.[1] Second, while the past literature has analyzed the relationship between patents and firm-level outcomes using firm-level aggregate data (Austin, 1993; Balasubramanian and Sivadasan, 2011; Farre-Mensa, Hegde and Ljungqvist, 2016; Kogan, Papanikolaou, Seru and Stoffman, 2013; Van Reenen, 1996), we will be able to investigate this question in our newly linked micro data tracking US firms and workers in Treasury tax files. These micro data allow us to decompose whether observed increases in worker compensation are driven by changes in the skill composition of workers at a firm, or driven by rent-sharing. More generally, this micro data will allow us to test more precisely for which sets of workers share in the rents from a new innovation (e.g. current workers versus future hires).

We start by documenting event study estimates around the time when a patent application is filed. Patent applications are preceded by stark runups in firm size – as measured by any of sales, capital, or employment. The fact that growing firms file patent applications is not surprising since in many cases the timing of the innovation underlying a given patent application likely precedes the year of application. However, this phenomenon has been obscured in some previous studies that condition on firm size in the years leading up to patent application (e.g., Balasubramanian and Sivadasan (2011)). Unlike firm size, measures of firm profitability (e.g. gross operating margins) are stable prior to the date of the patent application, but rise dramatically after a patent application is filed. Average wages per worker rise in advance of the date of the patent application, but stabilize once the application is submitted. To remove the influence of confounding pre-trends, we stratify our analysis by a measure of the ex ante value of the patent application. We find that wage growth, firm size measures, and measures of firm profitability all rise in the wake of a valuable patent application relative to applications that are ex ante less valuable. Taken together, these estimates suggest a causal relationship between the underlying innovation which gives rise to the patent application and firm and worker outcomes.

We then explore the impact of patent grants – as opposed to the innovation underlying the patent application – by instrumenting for patent grants with the leniency of the assigned patent examiner. We first document in our sample that the assigned USPTO examiner is strongly predictive of patent grants, and our preliminary estimates support this cross-examiner variation as a valid instrument: for example, the assigned examiner's grant rate is not correlated with lagged outcomes of the firm to which the patent is assigned. While our preliminary instrumental

[1]This variation was initially developed in Sampat and Williams (2014). The methodology is analogous to similar applications in other contexts, including investigations of the effects of incarceration length using variation across judges (Kling, 2006), the effects of foster care using variation across foster care case workers (Doyle, 2007, 2008) and investigations of disability insurance work incentives using variation across disability insurance examiners (Maestas, Mullen and Strand, 2010).

variable estimates are statistically imprecise, they suggest that patent grants have little effect on sales revenue or employment, but do lead to a mild increase in firm profitability and worker compensation.[2]

In both our event study estimates and our instrumental variables estimates, the observed increases in worker compensation could reflect either or both of two mechanisms: wages could rise because firms are sharing rents with workers, or worker compensation could be changing as firms change the skill mix of their employees (in either direction, skill upgrading or skill downgrading). In future work – in the near term – we will be able to test for each of these mechanisms by analyzing the effects of patent grants on different groups of employees. For example, by tracking the compensation of employees who were at the company prior to the patent grant longitudinally in workers' W2 Forms, we can test for rent sharing by investigating whether these individuals exhibit excess wage growth in the years following the patent grant relative to prior periods.[3] This approach mirrors the recent empirical literature on rent sharing (Guiso, Pistaferri and Schivardi, 2005; Guertzgen, 2009; Carlsson, Messina and Skans, 2011; Card, Cardoso and Kline, 2014) but relies on the timing of patent grants – and the patent examiner instrument – for identification, whereas the past literature has relied on unverifiable restrictions on the stochastic process underlying firm productivity. Separately, we can empirically assess whether firms engage in skill upgrading in response to being awarded patents by directly estimating how the education, experience, and "unobserved ability" of new hires changes in response to patenting, where unobserved ability is proxied by statistical "person fixed-effects" as in Abowd, Kramarz and Margolis (1999) and Card, Heining and Kline (2013). Though skill upgrading is one possibility, patenting may alternatively enable the creation of more routinized jobs that induce skill downgrading (Autor, Levy and Murnane, 2003).

Our rent sharing estimates will quantify the extent to which firms are obliged to share patent-generated rents with non-innovative employees. From a policy perspective, these estimates provide a useful input into assessing the economic incidence and efficiency of tax policies that impact research investments, and also shape our understanding of the distributional impacts of other policies affecting business profits, such as business tax reform. A sizable elasticity such as that documented by Van Reenen (1996) would suggest that research tax credits are captured, in part, by non-innovative workers. From a tax policy perspective, our focus on small firms – which have been excluded from prior studies, with the exception of the sample of small manufacturing firms covered in Balasubramanian and Sivadasan (2011) and the recent work of Farre-Mensa et al. (2016) – is of particular interest given that small firms often fail to benefit from the non-refundable research and experimentation (R&E) tax credit. Understanding the innovative contribution of small firms speaks to the value of potential reforms to the R&E tax

[2] We are working on improving the statistical precision of these estimates, and expect to make progress on that in the near-term.

[3] In addition to the prior firm-level studies cited above (Austin, 1993; Balasubramanian and Sivadasan, 2011; Farre-Mensa, Hegde and Ljungqvist, 2016; Kogan, Papanikolaou, Seru and Stoffman, 2013; Van Reenen, 1996), our work also relates to a recent study of Finnish employer-employee data (Toivanen and Vaananen, forthcoming) which documents some evidence that inventors receive a wage premia after being granted a patent.

credit and other policies that aim to increase private research investments. Finally, our estimates of the magnitude and timing of patent-generated rents are relevant to the design of "patent box" proposals that seek to levy taxes on profits derived from patented products.

Section 1 describes our data, and Section 2 provides some basic descriptive statistics. Section 3 outlines a simple theoretical framework that guides the interpretation of our empirical estimates. In Sections 4 and 5 we document our event study and instrumental variables estimates, respectively.

1 Data

Our empirical analysis takes advantage of a newly developed linkage of several databases, which provides us with panel data on the patent filings, patent grants, and outcomes of US firms.[4]

1.1 USPTO records of patent applications and granted patents

The core of our patent data is public-use administrative data on the universe of patent applications submitted to the US Patent and Trademark Office (USPTO) which allow us to observe the full set of US patent applications filed since late 2000.[5] After linking these US patent applications with data on patents granted to these applications, we match these data to the public-use USPTO PAIR (Patent Application Information Retrieval) administrative data, which records the key variables needed to construct our examiner leniency instrument – namely, the name of the USPTO examiner assigned to review the application and the variables needed for conditional random assignment (filing year and the so-called art unit and technology classes to which the patent application was assigned). Finally, because published patent applications are not required to list the assignee (owner) of the patent – which is needed in order to implement our firm name-based linkage between patent applications and the US Treasury tax files described below – we worked with the USPTO to gain access to a separate data file which allowed us to fill in missing assignee names for most (but not all) of the published patent applications.

1.2 US Treasury tax files

All variables are constructed from annual business tax returns over the years 1997-2014: C-Corporations (Form 1120), S-Corporations (Form 1120-S), and Partnerships (Form 1065).

[4] Appendix A describes our data construction in more detail.

[5] While these data on - successful and unsuccessful - patent applications are just starting to be analyzed, data on US granted patents has been heavily used and is available for earlier decades.

1.3 Linkage procedure

We built on the name standardization routine used by the National Bureau of Economic Research (NBER)'s Patent Data Project to accomplish a firm name-based merge of patent assignees to firm names. Specifically, we standardized the firm names in both the patent data and (separately) the business tax filings in order to infer that e.g. "ALCATEL-LUCENT U.S.A., INC..", "ALCATEL-LUCENT USA, INCORPORATED," and "ALCATEL-LUCENT USA INC" are all in fact the same firm. We then conducted a fuzzy merge of standardized patent names to standardized tax names using the SoftTFIDF algorithm based on a Jaro-Winkler distance measure.

Of course, a natural concern arises about the accuracy of a fuzzy merge of this type. To assess the quality of our merge, we have completed two preliminary quality checks – validation against a hand-coded sample, and validation against an inventor-based linkage of patent applications to individual tax returns developed by Bell et al. (2016).[6] These results of these two validation exercises are described in Appendix A, but in brief, the results of both data quality checks were quite reassuring.

1.4 Measurement of patent value

Because patent applications vary tremendously in their value (see, e.g., Pakes (1986)), we also collected data aiming to quantify the value of patent applications in order to explore heterogeneity across patent grants in the private value they generate for firms.

A variety of metrics have been proposed as measures of the value of granted patents. Dating back at least to the work of Trajtenberg (1990), a large literature has used forward patent citations as a measure of the value of granted patents. The key idea is that a given patent cites patents that it "builds on," so that more valuable patents should receive more forward citations. Other proposed measures for the value of granted patents have been developed using data on patent renewal behavior (Pakes, 1986; Schankerman and Pakes, 1986; Bessen, 2008), patent ownership reassignments (Serrano, 2010), patent litigation (Harhoff, Scherer and Vopel, 2003), and excess stock return values (Kogan, Papanikolaou, Seru and Stoffman, 2013). While widely used, these value measures encounter three challenges in our application. First, these measures are only defined for granted patents, whereas we would like to take advantage of data on patent applications (both successful and unsuccessful) and granted patents. Second, some (although not all) of these measures arguably correspond more closely to a measure of social value - or social spillovers, in the sense of social value minus private value - whereas in our context we are more interested in measuring firms' private value of a patent, given our interest in patent grants as a shifter of firm rents. This issue arises most sharply with forward citations, which are typically used as a measure of spillovers as in, e.g., Bloom, Schankerman and Van Reenen (2013). Third, all of these measures are defined ex post: citations,

[6] The first exercise – validation against a hand-coded sample – was done applying our linkage procedure to Compustat data, for disclosure reasons.

renewals, reassignments, and litigation are often measured many years after the initial patent award. But in our context - as in Kogan et al. (2013) - what is arguably more relevant is the expected private value of the patent at the time of the patent application or patent grant.

Motivated by these challenges, we instead rely on patent family size as our primary measure of patent value, as developed in Jonathan Putnam's dissertation (Putnam, 1996). Generally stated, a patent "family" is defined as a set of patent applications filed with different patenting authorities (e.g. US, Europe, Japan) that refer to the same invention. The key idea is that if there is a per-country cost of filing for a patent, firms will be more likely to file a patent application in multiple countries if they perceive the patent to have higher private value. Past work - starting with Putnam (1996) - has documented evidence that patent family size is correlated with other measures of patent value. We use patent family size as defined by Thompson Reuters' Innovation database, which counts the number of countries in which the patent application was filed with exactly the same priority.[7]

1.5 Measurement of rents and profits

As described in Card et al. (2016), empirical rent sharing estimates are often sensitive to a number of measurement issues, the most prominent of which is the choice of rent measure. The US Treasury tax files we use offer us the opportunity to explore this issue in significantly more detail than was possible in previous studies. We define the value added over which workers and firms bargain as sales minus the cost of goods sold plus (in manufacturing and mining) labor costs. In future work, we plan to experiment with different ways of scaling this measure including dividing by the total number of employees and using a quality adjusted measure that takes into account the occupational mix of the employees. Following recent work by Yagan (2015), we define business operating profits in the US Treasury tax files as sales minus cost of goods sold minus all deductions other than interest expenses, depreciation, and the domestic production activities deduction.[8] Since profits can be negative we use profit margins (as a fraction of sales) as a key measure of firm performance.

2 Descriptive statistics

Table 1 shows descriptive statistics for three sets of firms. Column 1 ("All Firms") contains all firm-year observations with tax returns in years 1997-2014. Column 2 ("Matched") is constructed by creating an annual observation for years 1997-2014 for each of the 83,455 firms that were successfully matched to at least one patent application.

[7] We have also investigated claims count as an alternative value measure that is fixed at the time of patent application, as proposed by Lanjouw and Schankerman (2001). The key idea here is that patents list "claims" over specific pieces of intellectual property, so that patents with more claims may be more valuable. Past work has documented mixed empirical evidence on whether that is a valid assumption based on correlations of claims counts with other value measures. Consistent with this mixed evidence, our preliminary analyses have suggested that claims count is much more weakly correlated with patent grants than is family size, so we here focus attention on family size.

[8] Yagan (2015) also adds back compensation to officers when defining profits. We do not because payments to capital owners represent profits after executives have been paid.

Column 3 ("Infrequent Patenters") restricts to the subset of firms in Column 2 which file 10 or fewer patent applications during our sample window. Given our focus on patents as a shifter of firm rents, a focus on firms that apply for at least one patent over our time period is natural. The motivation for the third ("infrequent patenters") sample is that many of our specifications will focus on this subset of firms which apply for patents sufficiently infrequently that the identity of their patent examiner is likely to be consequential.

Firms in our matched sample are much larger than firms in the full sample: for example, the median sales of firms in the full sample (Column 1) is around $230,000, whereas in our matched sample (Column 2) median sales is around $2.3 million. The average firm in our matched sample applies for roughly one patent a year, and receives roughly one patent every two years. The share of firms in the manufacturing sector in our matched sample is similar to the share in the full sample (around 5 percent).

Restricting to "infrequent patenters" in Column 3 narrows attention to 76,737 firms, which are roughly similar to the full sample of matched firms (Column 2) in terms of firm characteristics and sectoral mix. Hence, we would expect our analysis of infrequently patenting firms to yield results that are broadly representative of the matched sample.

3 Theoretical framework

It is useful to have a simple theoretical framework in mind when interpreting the possible effects of innovation and patent grants on firm behavior. We here present a variant of the model developed in Card et al. (2016) where firms have market power in both product and input markets. A firm produces output (Y) with a constant returns to scale production technology $f(.,.)$ that takes as input two types of labor (H and L) and capital (K):

$$\begin{aligned} Y &= Tf((1-\theta)L+\theta H, K) \\ &= Tf(N,K), \end{aligned}$$

where T is "physical" productivity, $\frac{\theta}{1-\theta}$ measures the relative efficiency of type H and L workers, and $N = (1-\theta)L+\theta H$ gives the total "efficiency units" of labor utilized by the firm.

The firm is a monopsonist that faces a constant supply elasticity of η_N of workers of either type. Capital is also supplied imperfectly with elasticity η_K. The firm's problem is to choose input prices to minimize costs:

$$\min_{w_L, w_H, r} w_L L(w_L) + w_H H(w_H) + rK(r) \ \ s.t.\ Tf(N,K) = Y$$

The first order conditions for this can be written:

$$\begin{aligned} w_L &= \frac{\eta_N}{1+\eta_N} T(1-\theta) f_1 \mu \\ w_H &= \frac{\eta_N}{1+\eta_N} T\theta f_1 \mu \\ r &= \frac{\eta_K}{1+\eta_K} T f_2 \mu, \end{aligned}$$

where μ is a Lagrange multiplier giving the marginal cost of output. Conditional on scale, wages are increasing in firm productivity with unit elasticity. However, worker's wages are "marked down" below their marginal revenue product in inverse relation to the elasticity of labor supply to the firm. As η_N approaches infinity, the labor market becomes competitive and wages are equated to marginal revenue products.

Product price depends on output according to the function $P(Y) = P_0 Y^{-1/\varepsilon}$, where $\varepsilon > 1$ gives the elasticity of product demand. Hence revenue ($R(Y)$) and marginal revenue can be written:

$$R(Y) = PY = P_0 Y^{1-1/\varepsilon}$$

$$R'(Y) = \left(1 - \frac{1}{\varepsilon}\right) P_0 Y^{-1/\varepsilon}.$$

The firm will equate marginal revenue $R'(Y)$ to marginal cost μ.

Imposing this equality, we can write the ratio of labor costs to revenue – the "compensation share" – as follows:

$$\begin{aligned} \frac{w_L L + w_H H}{R} &= \frac{\frac{\eta_N}{1+\eta_N} T f_1 N \left(1 - \frac{1}{\varepsilon}\right) P_0 Y^{-1/\varepsilon}}{P_0 Y^{1-1/\varepsilon}} \\ &= \frac{\varepsilon - 1}{\varepsilon} \frac{\eta_N}{1+\eta_N} \frac{T f_1 N}{Y} \\ &= \frac{\varepsilon - 1}{\varepsilon} \frac{\eta_N}{1+\eta_N} \lambda\left(\frac{K}{N}\right) \end{aligned}$$

where $\lambda\left(\frac{K}{N}\right)$ gives the ratio of the marginal product of labor to output. Note that if $f(.,.)$ were Cobb-Douglas this last component would be a technological constant.

The above expression shows that compensation share is the product of three components. The first component involves the product market demand elasticity ε. An increase in ε increases the compensation margin since workers capture only a fraction of the rents generated by the firm's market power. A patent grant can be thought of as a decrease in ε which should lead to a decrease in the compensation share. Second, the compensation share depends on the elasticity of labor supply to the firm. A higher elasticity of supply is analogous to workers have more bargaining power and raises the compensation margin. Finally, the third component reflects imperfect substitution between labor and capital.

A similar argument reveals that the firm's capital share can be written:

$$\frac{rK}{R} = \frac{\varepsilon - 1}{\varepsilon} \frac{\eta_K}{1+\eta_K} \left[1 - \lambda \left(\frac{K}{N} \right) \right]$$

which is increasing in ε and η_K. Likewise, the profit margin of the firm can be written:

$$1 - \frac{w_L L + w_H H + rK}{R} = 1 - \frac{\varepsilon - 1}{\varepsilon} \left[\frac{\eta_N}{1+\eta_N} \lambda \left(\frac{K}{N} \right) + \frac{\eta_K}{1+\eta_K} \left[1 - \lambda \left(\frac{K}{N} \right) \right] \right]$$

which is decreasing in ε, η_N, and η_K.

Hence, we expect a patent grant – which reduces ε – to reduce the compensation share and increase the firm's profit margin. Innovation itself has an unclear effect on the compensation share that depends on how it influences the capital-labor ratio and the production process itself. For example, an increase in physical productivity (T) that holds $f(.,.)$ constant should have no effect on the compensation share, capital share, or profit margin, but should raise w_L,w_H, r, and Y. However, it seems likely that an important part of what innovation does is to provide the firm with some temporary market power even before a patent is granted, e.g., due to product differentiation. To the extent that this occurs, innovations will have effects similar to patent grants.

4 Event study estimates

4.1 First applications

We begin with an analysis of the typical path of firm outcomes preceding and following a firm's first patent application. We use a standard event study specification of the form:

$$Y_{it} = \alpha_i + \gamma_{g(i)t} + \sum_{k \in \{-4,-3,-2,0,1,2,3,4,5\}} \beta_k D_{it}^k + \underline{\beta} \sum_{k<-4} D_{it}^k + \overline{\beta} \sum_{k>5} D_{it}^k + \varepsilon_{it} \tag{1}$$

where D_{it}^k is an indicator for firm i having first applied for a patent k periods in the past, α_i is a firm fixed effect, and $\gamma_{g(i)t}$ is a fixed effect for (2-digit industry)-by-(tax year) – that is, a flexible parametrization of time fixed effects. The coefficients β_k give the adjusted time path of mean outcomes relative to the period before first patent application (which has been normalized to zero). Here we have dealt with imbalance issues by "binning" periods greater than 5 or less than -4, which is reflected in the $\bar{\beta}$ and $\underline{\beta}$ coefficients that are assumed to be stable within end point bins. To address serial correlation in ε_{it}, all standard errors are clustered by firm.

We fit specifications of this type on our sample of infrequently patenting firms. The estimates for three firm size variables – log revenue, log employment, and log capital (defined as depreciable assets) – are plotted in

Figure 1(a). These estimates suggest that firms tend to grow abnormally fast in the years prior to submitting a patent application. This growth tapers off after the application is actually submitted. It is difficult to make causal inferences from this evidence because the growth trend might have tapered off even in the absence of the patent application. Balasubramanian and Sivadasan (2011) addressed this type of pattern by matching control firms to their patent applicants based on firm size in the year prior to the application. We are reticent to adopt this strategy because firm size is also an outcome of interest and so conditioning on lags may generate endogeneity biases.

We next fit specifications that interact the first-application event indicator variables with three ex-ante patent application quality bins – namely, the number of countries in which the application was filed: one country (that is, patent applications filed in the US only), 2-4 countries, or more than 5 countries. These interactions between first patent application indicator variables and ex-ante quality bins enable us to estimate quality-bin specific coefficients β_k^j where j indexes the quality bin.

$$Y_{it} = \alpha_i + \gamma_{g(i)t} + \sum_{k\in\{-4,-3,-2,0,1,2,3,4,5\}} \beta_k^j D_{it}^k \times 1(j(i)=j) \\ + \underline{\beta}^j \times 1(j(i)=j) \sum_{k<-4} D_{it}^k + \overline{\beta^j} \times 1(j(i)=j) \sum_{k>5} D_{it}^k + \varepsilon_{it} \tag{2}$$

Figure 1(b) plots the time path of one of our firm size variables (log revenue) relative to the date of the first patent application, separately for these three bins based on our ex ante measure of the value of a patent application. Prior to the date a patent application is filed, sales in these three sets of firms follow a strikingly similar trend. However, after the patent applications are file, firms filing higher quality patent applications exhibit greater revenue growth. Comparing across these three groups, this evidence suggests that innovations that are more privately valuable are associated with greater increases in firm size. However, the slow divergence in trajectories across the three sets of firms in Figure 1(b) makes it difficult to interpret whether these trends reflect differences in the value generated by the innovation itself, or value generated by a (lagged) patent grant.

Figure 2(a) reports event study coefficients for three measures of profitability: gross value added margins (defined as value added divided by revenue), gross earnings margins (defined as earnings before interest taxes and depreciation [EBITD] divided by revenue), and gross profit margins (defined as operating profits divided by sales). In contrast to the firm size variables plotted in Figure 1, these measures of profit margins are quite stable prior to the date of the first patent application. Gross earnings and profit margins then rise dramatically after a patent application is filed, suggesting that the underlying innovation leads the firm to become more profitable. Notably, value added as a fraction of revenue is remarkably stable both prior to and after the patent application is submitted, which is inconsistent with the notion that innovation manifests as a pure shock to "physical" total factor productivity – instead, this pattern suggests that these impacts are driven by changes in product market

power associated either with the innovation itself or the resulting patent grant. Figure 2(b) plots the patterns in gross earnings margins separately by our ex ante measure of patent application quality, which are somewhat noisy but the point estimates are consistent with slightly larger increases for patent applications that are filed in multiple countries.

Figure 3(a) reports event study coefficients for three measures of worker compensation. Average wages per worker rise before a patent application is filed and then stabilize after the date of first patent application. As with firm size, it is difficult to infer causal effects in the presence of such pre-trends. However Figure 3(b) shows that wages flatten out more in the wake of filing what are ex-ante less valuable patents. Indeed, the point estimate for the highest quality group three years after application is significantly larger than the corresponding point estimate for the lowest quality group. This pattern suggests that, on net, valuable innovations raise average wages. Figure 3(a) also decomposes the gross operating margin result of Figure 2(a) by treating as the dependent variable the ratio of labor compensation to revenue, which we term the "compensation margin." We also construct a version of this compensation variable which subtracts off officer revenue. Both of these compensation margin variables are quite stable prior to the date a patent application is filed, and both decline dramatically once the patent application is submitted. This is in keeping with our theoretical prediction that the compensation share should fall with an increase in market power (a decrease in ε). Appendix Figure C.1 further decomposes the compensation result into wage and non-wage components, and documents that the strongest response by far comes from salaries and wages (as opposed to e.g. benefits or pensions). This finding suggests that firms put downward pressure on wages in anticipation of a patent grant, though it is also possible that some of this effect results from compositional changes in the workforce – a possibility we will investigate more carefully in future work.

4.2 Multiple applications

Most firms in our sample apply more than once for a patent. To incorporate the information in subsequent grants, we estimate a two-sided distributed specification of the form:

$$Y_{it} = \alpha_i + \gamma_{g(i)t} + \sum_{k \in \{-4,-3,-2,0,1,2,3,4,5\}} \beta_k A_{it-k} + \underline{\beta} \sum_{k<-4} A_{it-k} + \bar{\beta} \sum_{k>5} A_{it-k} + \varepsilon_{it} \tag{3}$$

where A_{it} is the number of applications in year t and $\gamma_{g(i)t}$ again represents a fixed effect for (2-digit industry)-by-(tax year). If second and subsequent grants have the same dynamic effects as first grants, then this approach should provide more efficient estimates of the same event study coefficients. The results obtained through this approach (pictured in Appendix Figures C.2-C.4) are indeed more precise, as expected, and also smaller in magnitude –

suggesting that the effects of first patent applications and subsequent patent applications may be quite different.

Table 2 summarizes these estimates by estimating long differences specifications of the form:

$$Y_{it} - Y_{it-4} = \tilde{\alpha}_i + \tilde{\gamma}_{g(i)t} + \tilde{\beta} A_{it-4} + \tilde{\varepsilon}_{it}.$$

The coefficient $\tilde{\beta}$ estimates the impact of an additional patent application on the four year growth rate in a given outcome variable.

5 Instrumental variables estimates

To separate the effects of an innovation from the effects of the temporary monopoly associated with a patent grant we instrument for patent grants with examiner leniency, assuming quasi-random assignment of patent examiners to patent applications.

5.1 Measuring examiner leniency

We estimate examiner leniency using all utility patent applications with non-missing examiner names in the USPTO data from 2001 to 2012, including those not merged to the business tax records. As suggested by Feng and Jaravel (2016) we discard all non-progenitor patent applications (that is, we discard e.g. continuation and divisional applications), which leaves us with roughly 1.63 million patent applications over our sample window. To explain our approach it will be useful to develop some notation. Define G_j^4 as an indicator for whether an application $j \in \{1,...,J\}$ is granted within four years of submission (roughly the median in our sample). Let $ART_j \in \{1,...,R\}$ denote the intersection of application j's "art unit" code with the year in which the application was submitted; we henceforth refer to this intersection as the "art stratum." Each stratum r has a collection $\mathscr{E}_r$ of patent examiners that we number from 1 to M_r. Let $EX_j \in \mathscr{E}_{ART_j}$ denote the identity of the patent examiner assigned to application j. Every application has a potential grant status $G_j^4(m)$ that would result if it were assigned to examiner m. Realized grant status is $G_j^4 = G_j^4(EX_j)$.

The nominal assumption in this literature is that applications are randomly assigned to examiners within art strata, which would imply that:

$$\left(G_j^4(1),...,G_j^4(M_r)\right) \perp EX_j | ART_j.$$

Hence, if we look within an art stratum r, differences in grant rates across two examiners m and m' reflect average causal effects:

$$E\left[G_j^4 | EX_j = m, ART_j = r\right] - E\left[G_j^4 | EX_j = m', ART_j = r\right] = E\left[G_j^4(m) - G_j^4\left(m'\right) | ART_j = r\right].$$

Consequently, one approach would be to simply use the difference between each examiner's mean allowance rate $\bar{G}_{rm}$ and the average allowance rate $\bar{G}_r$ in that art stratum as a measure of leniency. However art strata are sometimes small, which suggests shrinking these estimated examiner effects towards zero could yield a much more powerful predictor of firm responses. For this reason, we apply a set of stratum-specific shrinkage weights $\{\hat{\pi}_r\}$ to these differences to obtain our improved leniency estimator $\hat{\ell}_{rm} = \hat{\pi}_r \left(\bar{G}_{rm} - \bar{G}_r\right)$ which, under weak assumptions, has an interpretation as a best linear predictor of true leniency (see Appendix B for details).

5.2 First stage estimates

To investigate whether there is a first stage relationship in our estimation sample, we estimate an analogue of the application event study in (3) taking the form:

$$G_{it} = \alpha_i + \gamma_{g(i)t} + \sum_{k\in\{-4,-3,-2,0,1,2,3,4,5\}} \beta_k \sum_{j=1}^{A_{it-k}} \hat{\ell}_{it-kj} + \underline{\beta} \sum_{k<-4} \sum_{j=1}^{A_{it-k}} \hat{\ell}_{it-kj} + \bar{\beta} \sum_{k>5} \sum_{j=1}^{A_{it-k}} \hat{\ell}_{it-kj} + \varepsilon_{it} \tag{4}$$

where $\left\{\hat{\ell}_{imj}\right\}_{j=1}^{A_{im}}$ denotes the set of examiner leniencies corresponding to firm i's applications submitted in year m. Hence, our specification consists of regressing the number of patents granted to firm i in year t (G_{it}) against leads and lags of total exposure to examiner leniency, firm fixed effects, and 2-digit industry-by calendar year fixed effects.

Figure 4(a) documents the results of this exercise in our sample of infrequently patenting firms. As expected, future leniency is unrelated to the number of grants received. But past leniency exposure is a strong positive predictor of grants in the four years following application.[9] Note that, according to our shrinkage model, we would expect the sum of the coefficients $\{\beta_k\}_{k=0}^{4}$ to equal one in the sample of all patents (since the examiner leniency measure is the best linear predictor of G_j^4). In fact, the sum is approximately 0.9, which suggests our shrinkage procedure is not heavily affected by restricting to the estimation sample of infrequent patentors. Figure 4(b) shows that higher quality applications (that is, applications filed in multiple countries) are more likely to yield grants which is consistent with family size being an objective proxy of patent quality.

5.3 Reduced form estimates

We now consider the effects of chance variation in examiner leniency on firm outcomes. To investigate this issue we again estimate specifications of the form given in (4) but with the dependent variable changed to firm outcomes.

[9] The negative coefficients in years 4 and 5 after application reflect that the additional grants in years 1-3 are being diverted from those later periods.

Our very preliminary estimates exploiting this quasi-experimental measure of examiner leniency are quite statistically imprecise, but currently tell a qualitatively different story about the effects of patent grants as opposed to applications.

Figure 5 plots the resulting estimates for our three firm size variables – log revenue, log employment, and log capital (defined as depreciable assets). We find no evidence that log sales depend on future examiner leniency, which corroborates the instrument's validity. However, the lags of examiner leniency are slightly negatively (though insignificantly) correlated with firm size, suggesting that patent grants may reduce revenue. Though somewhat surprising, this result is consistent with the hypothesis that the temporary monopoly power associated with patent grants leads to a decrease in output. More surprising are the results of Figure 6 which show only weak evidence of a (somewhat delayed) increase in profit margins.

Figure 7 shows corresponding estimates for worker compensation. Once again, we find no correlation between leads of examiner leniency and the labor compensation ratio. However, lags of examiner leniency are weakly positively correlated with average wages with a time pattern similar to firm profit margins (see also Table 3). This suggests that once a patent is actually granted, workers are able to capture some of the ensuing product market rents. Going forward, we are working on improving our measure of examiner leniency in order to obtain more statistically precise estimates.

6 Conclusion

To be added.

References

Abowd, John, Francis Kramarz, and David Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

Austin, David, "An event-study approach to measuring innovative output: The case of biotechnology," *American Economic Review Papers & Proceedings*, 1993, *83* (2), 253–258.

Autor, David, Frank Levy, and Richard Murnane, "The skill content of recent technological change: An empirical exploration," *Quarterly Journal of Economics*, 2003, *116* (4), 1279–1334.

Balasubramanian, Natarajan and Jagadeesh Sivadasan, "What happens when firms patent? New evidence from U.S. economic census data," *Review of Economics and Statistics*, 2011, *93* (1), 126–146.

Barth, Erling, Alex Bryson, James Davis, and Richard Freeman, "The contribution of dispersion across plants to the increase in US earnings dispersion," 2011. Institute for Social Research University of Oslo Working Paper.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen, "The lifecycle of inventors," 2016. Stanford Working Paper.

Bessen, James, "The value of U.S. patents by owner and patent characteristics," *Research Policy*, 2008, *37* (5), 932–945.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, "Identifying technology spillovers and product market rivalry," *Econometrica*, 2013, *81*, 1347–1393.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining and the gender wage gap: A direct assessment," 2014. UC-Berkeley Working Paper.

_ , _ , Joerg Heining, and Patrick Kline, "Firms and labor market inequality: Evidence and some theory," 2016. UC-Berkeley Working Paper.

_ , Joerg Heining, and Patrick Kline, "Workplace heterogeneity and the rise of West German wage inequality," *Quarterly Journal of Economics*, 2013, *128* (3), 967–1015.

Carlsson, Mikael, Julian Messina, and Oskar Nordstrom Skans, "Wage adjustment and productivity shocks," 2011. Sveriges Riksbank Working Paper Series No. 253.

Doyle, Joseph J., "Child protection and child outcomes: Measuring the effects of foster care," *American Economic Review*, 2007, *97* (5), 1583–1610.

_ , "Child protection and adult crime: Using investigator assignment to estimate causal effects of foster care," *Journal of Political Economy*, 2008, *116* (4), 746–770.

Farre-Mensa, Joan, Deepak Hegde, and Alexander Ljungqvist, "The bright side of patents," USPTO Economic Working Paper Series 2016.

Feigenbaum, James, "Automated Census Record Linking," Technical Report, Harvard University 2015.

Feng, Josh and Xavier Jaravel, "Patent trolls and the patent examination process," 2016. Harvard Working Paper.

Flinn, Christopher, "Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates," *Econometrica*, 2006, *74* (4), 1013–1062.

_ , *The Minimum Wage and Labor Market Outcomes*, MIT Press, 2010.

Guertzgen, Nicole, "Rent-sharing and collective bargaining coverage: Evidence from linked employer-employee data," *Scandinavian Journal of Economics*, 2009, *111* (2), 323–349.

Guiso, Luigi, Luigi Pistaferri, and Fabiano Schivardi, "Insurance within the firm," *Journal of Political Economy*, 2005, *113* (5), 1054–1087.

Hagedorn, Marcus and Iourii Manovski, "The cyclical behavior of equilibrium unemployment and vacancies revisited," *American Economic Review*, 2008, *98* (4), 1692–1706.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "The NBER patent citations data file: Lessons, insights, and methodological tools," 2001. NBER working paper #8498.

_ , _ , and _ , "Market value and patent citations," *RAND Journal of Economics*, 2005, *36* (1), 16–38.

Harhoff, Dietmar, F.M. Scherer, and Katrin Vopel, "Citations, family size, opposition, and the value of patent rights," *Research Policy*, 2003, *32* (8), 1343–1363.

Kling, Jeffrey, "Incarceration length, employment, and earnings," *American Economic Review*, 2006, *96* (3), 863–876.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological innovation, resource allocation, and growth," 2013. MIT mimeo.

Lanjouw, Jean and Mark Schankerman, "Characteristics of patent litigation: A window on competition," *RAND Journal of Economics*, 2001, *32* (1), 129–151.

Maestas, Nicole, Kathleen Mullen, and Alexander Strand, "Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt," 2010. University of Michigan Retirement Research Center working paper WP 2010-241.

Pakes, Ariel, "Patents as options: Some estimates of the value of holding European patent stocks," *Econometrica*, 1986, *54* (4), 755–784.

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva, "Optimal taxation of top labor incomes: A tale of three elasticities," *American Economic Journal: Economic Policy*, 2014, *6* (1), 230–271.

Putnam, Jonathan, "The value of international patent protection," 1996. Yale PhD dissertation.

Sampat, Bhaven and Heidi Williams, "How do patents affect follow-on innovation? Evidence from the assignment of patent applications to examiners," 2014. MIT mimeo.

Schankerman, Mark and Ariel Pakes, "Estimates of the value of patent rights in European countries during the post-1950 period," *Economic Journal*, 1986, *96* (384), 1052–1076.

Serrano, Carlos, "The dynamics of the transfer and renewal of patents," *RAND Journal of Economics*, 2010, *41* (4), 686–708.

Shimer, Robert, "The cyclical behavior of equilibrium unemployment and vacancies," *American Economic Review*, 2005, *95* (1), 25–49.

Toivanen, Otto and Lotta Vaananen, "Returns to inventors," forthcoming. Review of Economics and Statistics.

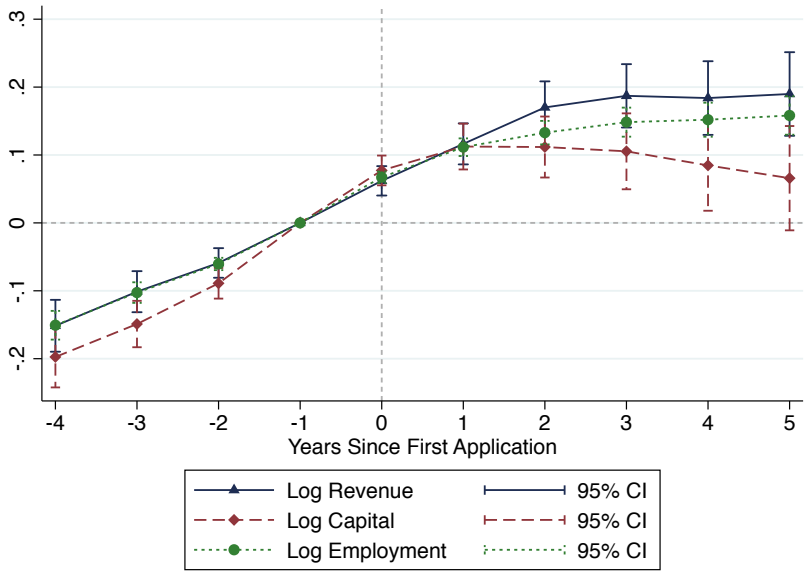
Trajtenberg, Manuel, *Economic Analysis of Product Innovation: The Case of CT Scanners*, Harvard University Press, 1990.

Van Reenen, John, "The creation and capture of rents: Wages and innovation in a panel of U.K. companies," *Quarterly Journal of Economics*, 1996, *111* (1), 195–226.

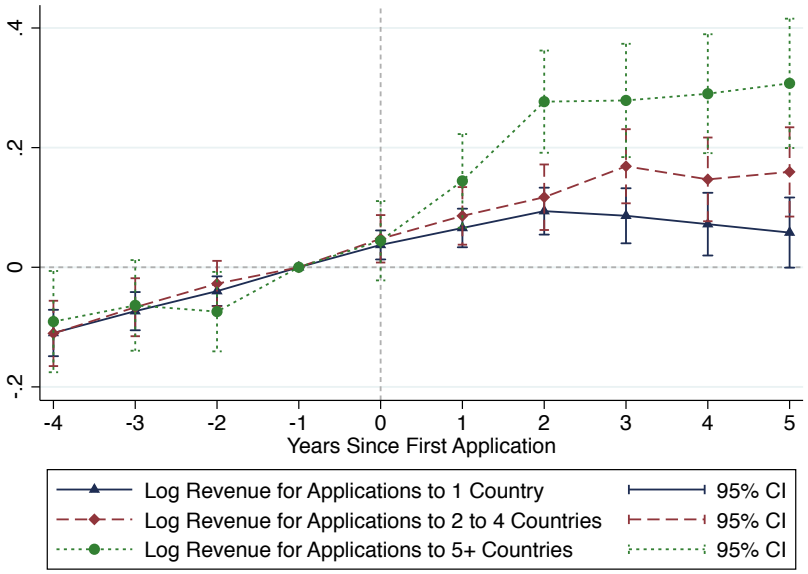
Yagan, Danny, "Capital Tax Reform and the Real Economy: The Effects of the 2003 Dividend Tax Cut," *American Economic Review*, 2015, *105* (12), 3531–3563.

Figure 1: Event Study of First Patent Applications on Firm Size

(a) Firm Size Measures



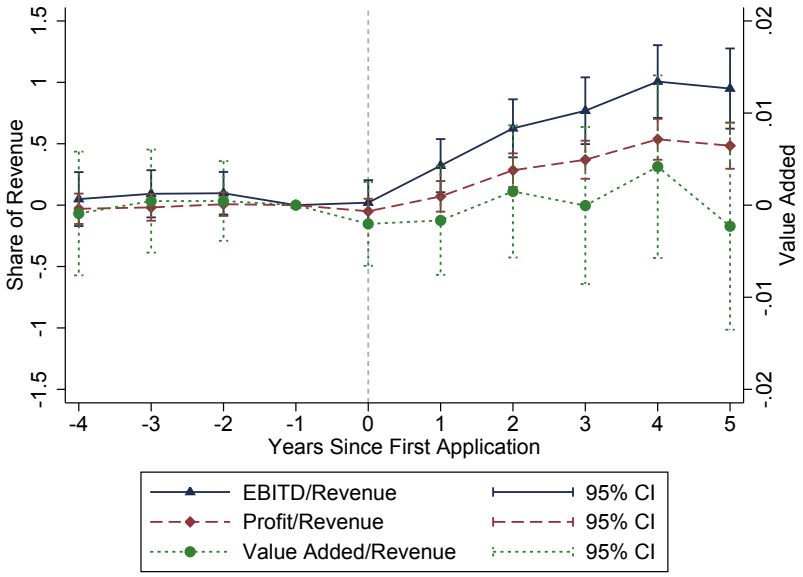
(b) Firm Size by Ex-Ante Patent Quality



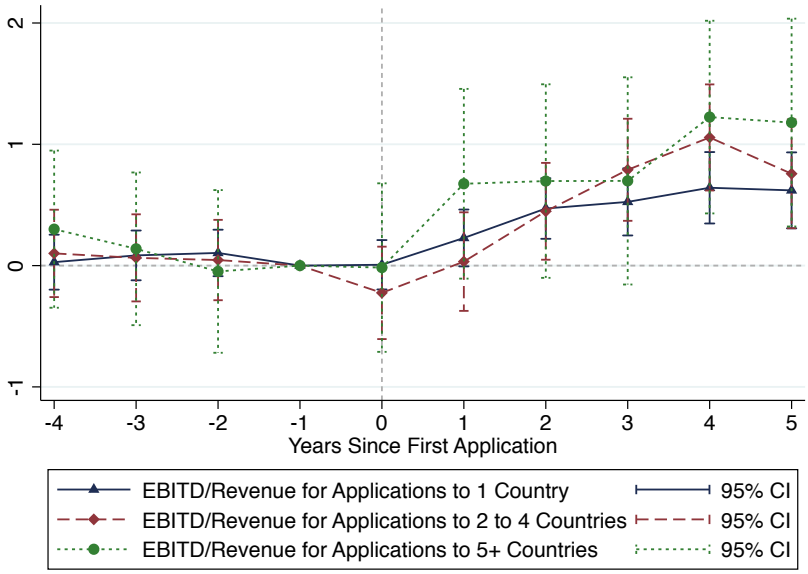
Note: These figures show estimates of β_k from the event study specification in equation 1. Confidence intervals based upon standard errors clustered on firm. Panel (a) presents three separate sets of estimates to show how measures of firm size evolve around the firm's first patent application from 2000 to 2010 among infrequently patenting firms, i.e., firms with 10 or fewer patent applications from 2000-2010. Panel (b) shows similar estimates of log revenue around all patent applications from 2000 to 2010 among infrequently patenting firms for three types of applications: patent applications sent to one country, to 2 to 4 other countries, and to more than 4 countries. In our sample of matched applications, 43.7 percent of applications are sent to one country, 41.9 are sent to 2 to 4 countiries, and 14.4 are sent to at least 5 countries. Firm size data come from US Treasury business tax files from 1997-2014. Revenue and capital are from annual business tax forms F1120, F1120S, and F1065 of active firms (i.e., firms with non-missing and non-zero total income and non-missing total deductions), and employment is the total number of employees associated with the firm in form W2 filings. Revenue and capital are measured in 2014 dollars. There are 225,832 and 197,998 firm-years for infreqeuntly patenting firms with positive revenue and capital, respectively. See Appendix A.2 and A.3 for additional details on patent and tax variable definitions.

Figure 2: Event Study of Patent Applications on Firm Performance

(a) Firm Performance Measures



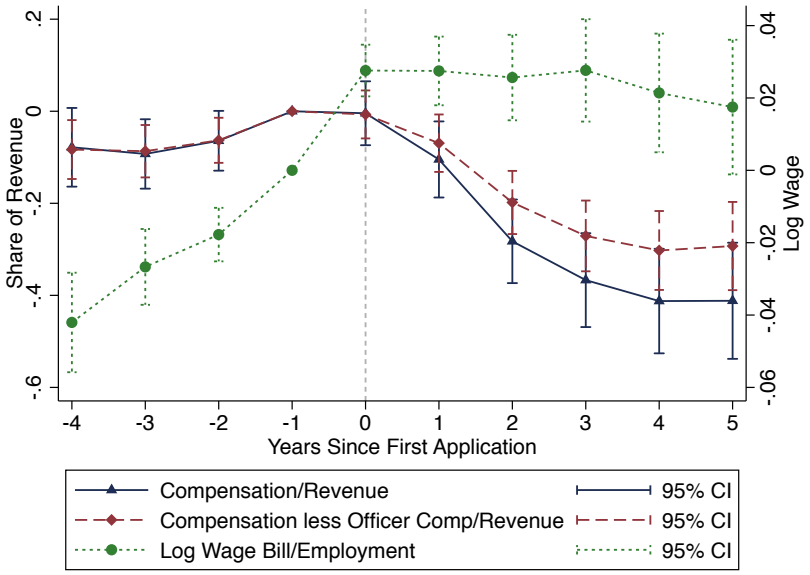
(b) Firm Performance by Ex-Ante Patent Quality



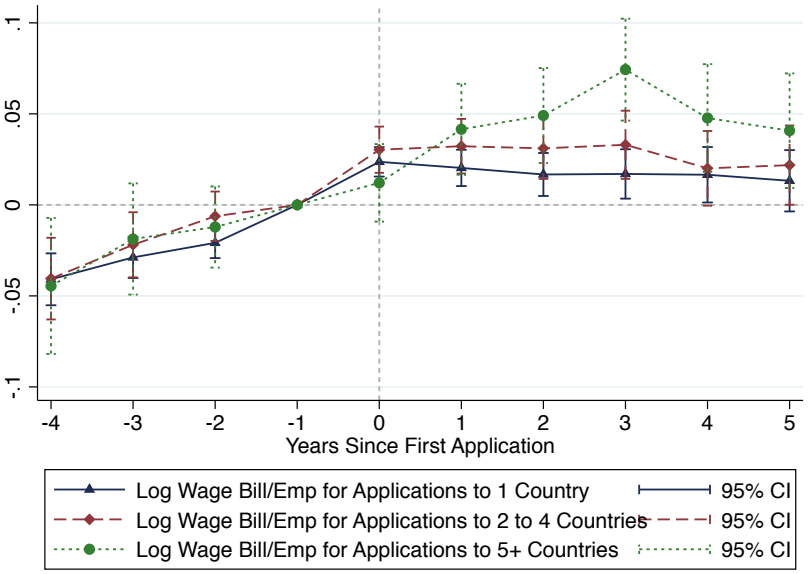
Note: These figures show estimates of β_k from the event study specification in equation 1. Confidence intervals based upon standard errors clustered on firm. Panel (a) presents three separate sets of estimates to show how measures of firm performance evolve around the firm's first patent application from 2000 to 2010 among infrequently patenting firms, i.e., firms with 10 or fewer patent applications from 2000-2010. Panel (b) shows similar estimates of EBITD margins for all patent applications from 2000 to 2010 among infrequently patenting firms for three types of applications: patent applications sent to one country, to 2 to 4 other countries, and to more than 4 countries. In our sample of matched applications, 43.7 percent of applications are sent to one country, 41.9 are sent to 2 to 4 countiries, and 14.4 are sent to at least 5 countries. Firm performance data are from annual business tax forms F1120, F1120S, and F1065 of active firms (i.e., firms with non-missing and non-zero total income and non-missing total deductions) over the period 1997-2014. Profits are operating revenues less costs of good sold and deductions other than interest expenses, depreciation, and the domestic production acticities deduction. See Appendix A.2 and A.3 for additional details on patent and tax variable definitions. There are 151,244, 130,519, and XXX firm-years for infrequently patenting firms for the EBITD, for the Profit, and Value Added estimates respectively.

Figure 3: Event Study of Patent Applications on Worker Outcomes

(a) Worker Outcomes



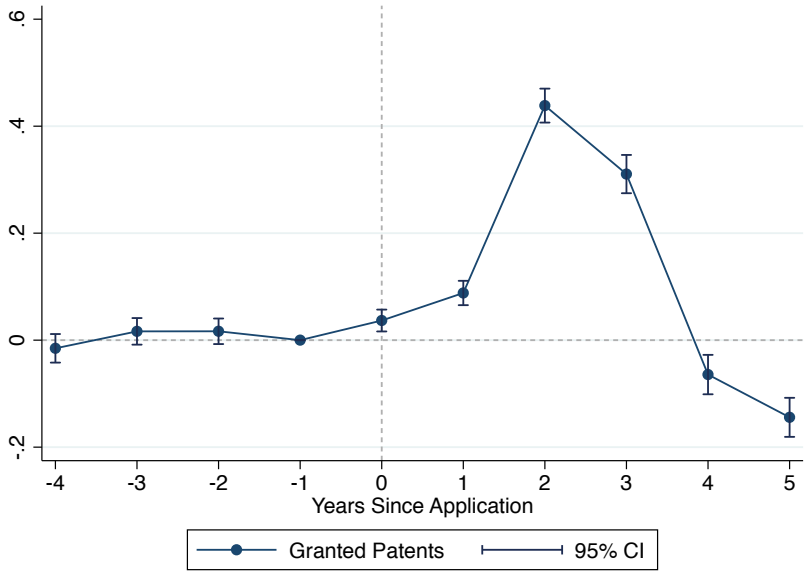
(b) Average Wages by Ex-Ante Patent Quality



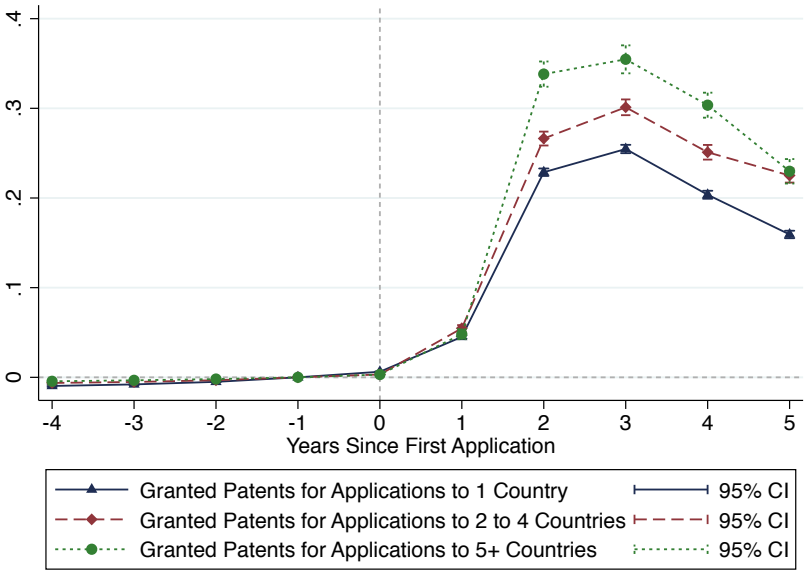
Note: These figures show estimates of β_k from the event study specification in equation 1. Confidence intervals based upon standard errors clustered on firm. Panel (a) presents three separate sets of estimates to show how worker outcomes evolve around the firm's first patent application from 2000 to 2010 among infrequently patenting firms, i.e., firms with 10 or fewer patent applications from 2000-2010. Panel (b) shows similar estimates of average worker compensation for all patent applications from 2000 to 2010 among infrequently patenting firms for three types of applications: patent applications sent to one country, to 2 to 4 other countries, and to more than 4 countries. In our sample of matched applications, 43.7 percent of applications are sent to one country, 41.9 are sent to 2 to 4 countiries, and 14.4 are sent to at least 5 countries. Revenue, compensation, and compensation less officer compensation are from annual business tax forms F1120, F1120S, and F1065 of active firms (i.e., firms with non-missing and non-zero total income and non-missing total deductions) over the period 1997-2014. The wagebill and employment data are totals of employees associated with the firm in form W2 filings. See Appendix A.2 and A.3 for additional details on patent and tax variable definitions. There are XXXX, XXXXX, and XXX firm-years for infrequently patenting firms for the compensation, net compensation, and average compensation estimates respectively. The wagebill is measured in 2014 dollars.

Figure 4: Event Study Impacts of Examiner Leniency on Patent Grants

(a) Granted Patents

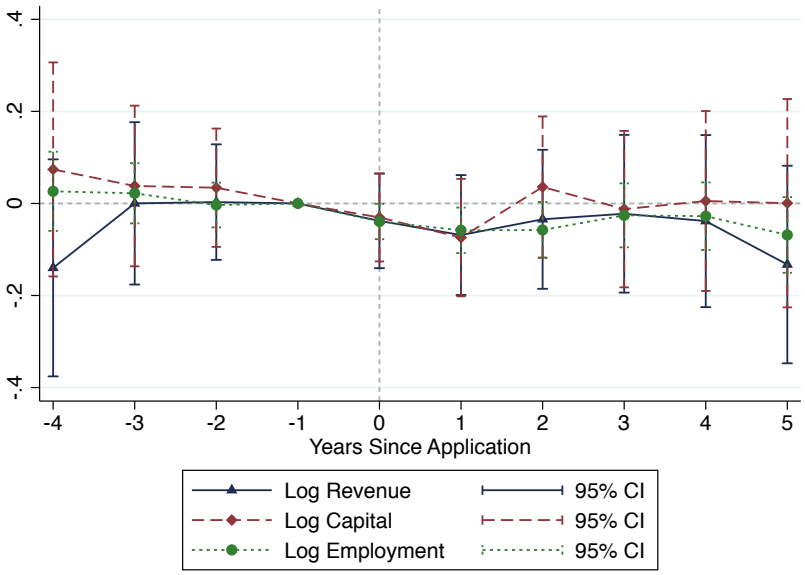


(b) Granted Patents by Ex-Ante Patent Quality



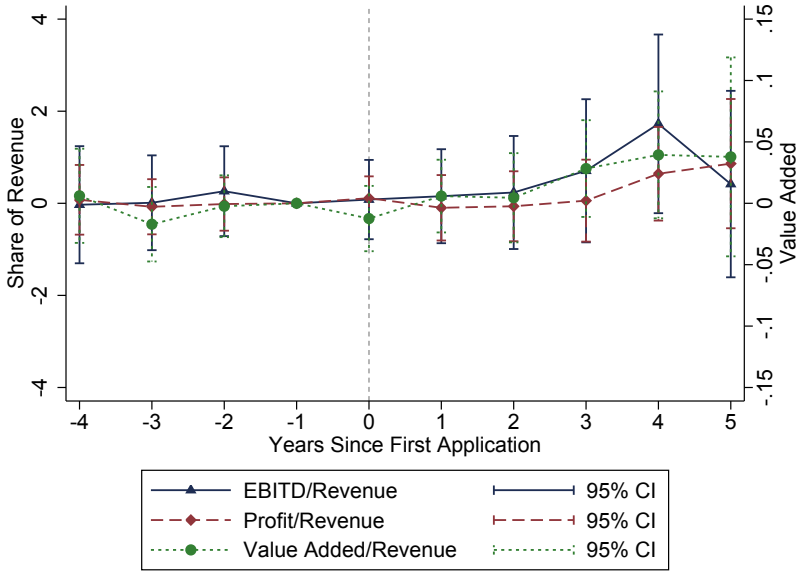
Note: These figures show estimates on the number of granted patents from event study specifications. Confidence intervals based upon standard errors clustered on firm. Panel (a) shows estimates of β_k from the event study specification in equation 2, and Panel (b) shows estimates from the specification in equation 1 [the equation 2 version is in progress]. See Appendix A.2 for additional details on patent variable definitions. The sample for both figures are infrequently patenting firms, i.e., firms with 10 or fewer patent applications from 2000-2010.

Figure 5: Event Study Impacts of Examiner Leniency on Firm Size



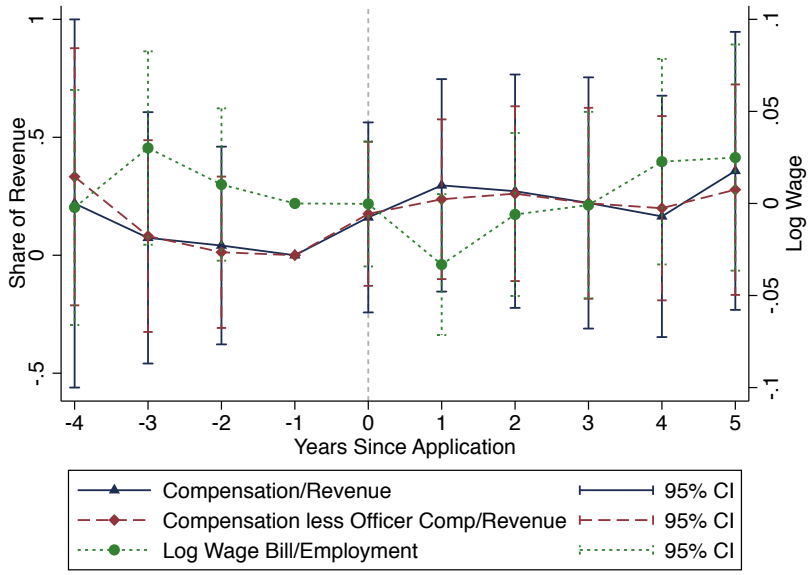
Note: This figure shows estimates of the effect of examiner leniency on three separate measures of firm size. Confidence intervals based upon standard errors clustered on firm. The sample is infrequently patenting firms, i.e., firms with 10 or fewer patent applications from 2000-2010. Firm size data come from US Treasury business tax files from 1997-2014. Revenue and capital are from annual business tax forms F1120, F1120S, and F1065 of active firms (i.e., firms with non-missing and non-zero total income and non-missing total deductions), and employment is the total number of employees associated with the firm in form W2 filings. Revenue and capital are measured in 2014 dollars.

Figure 6: Event Study Impacts of Examiner Leniency on Firm Performance



Note: This figure shows estimates of the effect of examiner leniency on three separate measures of firm performance. Confidence intervals based upon standard errors clustered on firm. The sample is infrequently patenting firms, i.e., firms with 10 or fewer patent applications from 2000-2010. Firm performance data are from annual business tax forms F1120, F1120S, and F1065 of active firms (i.e., firms with non-missing and non-zero total income and non-missing total deductions) over the period 1997-2014. Profits are operating revenues less costs of good sold and deductions other than interest expenses, depreciation, and the domestic production acticities deduction. See Appendix A.2 and A.3 for additional details on patent and tax variable definitions.

Figure 7: Event Study Impacts of Examiner Leniency on Worker Outcomes



Note: This figure shows estimates of the effect of examiner leniency on three separate measures of worker outcomes. Confidence intervals based upon standard errors clustered on firm. The sample is infrequently patenting firms, i.e., firms with 10 or fewer patent applications from 2000-2010. Revenue, compensation, and compensation less officer compensation are from annual business tax forms F1120, F1120S, and F1065 of active firms (i.e., firms with non-missing and non-zero total income and non-missing total deductions) over the period 1997-2014. The wagebill and employment data are totals of employees associated with the firm in form W2 filings. See Appendix A.2 and A.3 for additional details on patent and tax variable definitions. The wagebill is measured in 2014 dollars.

Table 1: Summary Statistics

	(1) All Firms	(2) Matched	(3) Infrequent Patenters
Medians [IQRs]			
Sales ($ 000s)	227.79	2333.39	2068.91
		[353.36,16769.72]	[327.5,12661.48]
Intermediate Input Costs ($ 000s)	8.59	920.07	785.97
		[30.33,8801.56]	[23.6,6581.54]
Value Added ($ 000s)	141.04	603.42	548.95
		[56.04,4545.52]	[51.49,3610.88]
Employment		18	17
		[0,75.95]	[0,67.46]
Labor Costs ($ 000s)	19.76	3588.96	2484.61
		[811.5,26957.62]	[682.75,12406.42]
Means (SDs)			
Log(Sales)	5.62	14.82	14.59
	(2.05)	(3)	(2.82)
Positive Input Costs	.54	.97	.97
	(.49)	(.17)	(.17)
Positive Value Added	.88	.99	.99
	(.32)	(.08)	(.08)
Log(Employment)		3.04	2.93
		(2.05)	(1.93)
Log(Labor Costs)	4.41	15.46	15.03
	(2.01)	(2.44)	(2.2)
Firm Age	11.38	0	0
	(11.67)	(0)	(0)
Manufacturing	.05	.06	.06
	(.21)	(.23)	(.23)
Scientific	.06	.02	.02
	(.23)	(.15)	(.15)
Pharmaceutical	0	0	0
	(.02)	(.04)	(.04)
# Applications		.95	.11
		(24.63)	(.42)
# Grants		.65	.06
		(16.36)	(.28)
Firms	5,126,047	83,455	76,737
Firm-Years		1,502,190	1,381,266

NOTES: This table reports statistics from firm-year panel data on innovative activity (number of applications and grants) and firm performance (financial and employment outcomes). Each column corresponds to different samples. The “All" sample contains all firm-year observations with tax returns in years 1997-2014 (based on the weighted IRS statistics of income sample). Annual measures of innovative activity and economic performance are then merged to this spine for each firm-year. The “Matched” sample is constructed by creating an annual observation for years 1997-2014 for each of the 83,455 firms that we successfully matched with a firm name on a patent application. “Infrequent patenters" are patenting firms with 10 or fewer patent applications. Numbers in brackets are [25th percentile, 75th percentile] and standard deviations are in parenthesis. “Percentiles” are presented as means of all observations in the *P*th percentile to protect taxpayer anonymity. Input costs are measured as sales less value added. Manufacturing is defined as any NAICS code beginning with “31", “32", or“33". Scientific is defined as any NAICS code beginning with “5415", “5416", or “5417". Pharmaceutical is defined as any NAICS code beginning with “3254". All monetary values are in 2014 dollars unless otherwise noted.

Table 2: Long Difference Effect of Prior Patent Applications

Outcome	Estimate	Std. Error	t-stat	N
Granted Patents	0.39	0.00	109.70	1,074,318
Log Revenue	-0.02	0.01	3.35	131,522
Log Capital	-0.07	0.01	8.75	112,984
Log Employment	-0.03	0.00	8.82	258,057
Value Added/Revenue	0.00	0.00	2.49	131,518
EBITD/Revenue	0.16	0.06	2.50	80,235
Profit/Revenue	0.12	0.04	3.08	69,654
EBITD/Capital	-0.13	0.14	0.92	79,399
Compensation/Revenue	-0.06	0.03	2.48	131,452
Log Wagebill/Employment	-0.01	0.00	3.79	257,965
Log Wagebill	-0.03	0.00	9.02	257,965

NOTES: This table shows estimates of $\tilde{\beta}_k$ from the long difference specification that regresses four year outcome growth $Y_{it} - Y_{it-4}$ on prior patent applications A_{it-4} and 2-digit industry-by-year fixed effects. Each row is a seperate regressions. Standard errors clustered on firm. The sample is infrequently patenting firms, i.e., firms with 10 or fewer patent applications from 2000-2010. Annual granted patents is a flow, but outcome growth is better measured as a change in the stock. We convert granted patents to a stock by summing the numner of granted patents over the past four years and use this sum as the outcome for the granted patents specification. Revenue, capital, value added, EBITD, profit, and compensation are from annual business tax forms F1120, F1120S, and F1065 of active firms (i.e., firms with non-missing and non-zero total income and non-missing total deductions) over the period 1997-2014. The wagebill and employment data are totals of employees associated with the firm in form W2 filings. See Appendix A.2 and A.3 for additional details on patent and tax variable definitions. The number of observations for granted patents is larger since we do not restrict the sample to active firms to allow for the possibility that firms have innovative activity before generating income. All monetary values are in 2014 dollars.

Table 3: Long Difference Effect of Prior Patent Examiner Leniency

Outcome	Estimate	Std. Error	t-stat	N
Granted Patents	0.90	0.02	38.61	1,074,318
Log Revenue	-0.02	0.06	0.34	131,522
Log Capital	-0.01	0.07	0.13	112,984
Log Employment	0.00	0.02	0.06	258,057
Value Added/Revenue	0.03	0.02	2.14	131,518
EBITD/Revenue	0.35	0.53	0.66	80,235
Profit/Revenue	0.28	0.30	0.95	69,654
EBITD/Capital	-0.69	1.17	0.59	79,399
Compensation/Revenue	0.11	0.19	0.57	131,452
Log Wagebill/Employment	0.03	0.02	1.41	257,965
Log Wagebill	0.03	0.03	1.13	257,965

NOTES: This table shows estimates of $\tilde{\beta}_k$ from the long difference specification that regresses four year outcome growth $Y_{it} - Y_{it-4}$ on lagged patent examiner leniency $\hat{l}_{it-4}$ and 2-digit industry-by-year fixed effects. Each row is a seperate regressions. Standard errors clustered on firm. The sample is infrequently patenting firms, i.e., firms with 10 or fewer patent applications from 2000-2010. Annual granted patents is a flow, but outcome growth is better measured as a change in the stock. We convert granted patents to a stock by summing the numner of granted patents over the past four years and use this sum as the outcome for the granted patents specification. Revenue, capital, value added, EBITD, profit, and compensation are from annual business tax forms F1120, F1120S, and F1065 of active firms (i.e., firms with non-missing and non-zero total income and non-missing total deductions) over the period 1997-2014. The wagebill and employment data are totals of employees associated with the firm in form W2 filings. See Appendix A.2 and A.3 for additional details on patent and tax variable definitions. The number of observations for granted patents is larger since we do not restrict the sample to active firms to allow for the possibility that firms have innovative activity before generating income. All monetary values are in 2014 dollars.

A Appendix: Data appendix

In this appendix, we describe our data construction in more detail.

A.1 Description of patent data

Our patent data build draws on several sources. Three identification numbers are relevant to have in mind with these datasets. First, publication numbers are unique identifiers assigned to published patent applications. Second, application numbers are unique identifiers assigned to patent applications that in practice are quite similar to publication numbers, but sometimes one application number is associated with multiple publication numbers. Finally, patent grant numbers are unique identifiers assigned to granted patents. Note that one patent application number can be granted more than one patent.

1. Census of published USPTO patent applications. We observe the census of published (accepted and rejected) patent applications published by the US Patent and Trademark Office (USPTO). Our original source for this data is a set of bulk XML files hosted by Google.[10] The underlying XML file formats were often inconsistent across years, so in the process of parsing these XML files to flat files we attempted to validate the data against outside sources wherever possible. These records are at the publication number level.

2. Census of granted USPTO patents. For the published USPTO patent applications in our data, we wish to observe which of those applications were granted patents. Our original source for this data is a set of bulk XML files hosted by Google.[11] As with the published USPTO patent applications data, the underlying XML file formats were often inconsistent across years, so in the process of parsing these XML files to flat files we attempted to validate the data against outside sources wherever possible. These records are at the patent grant number level.

3. USPTO patent assignment records. Some published USPTO patent applications are missing information on assignee names. Based on informal conversations with some individuals at the USPTO, we learned that the USPTO patent assignment database offers a way to fill in many – although not all – of these missing values, by taking advantage of the fact that assignee names are recorded in this separate database. Specifically, the patent assignment records include both initial assignments and re-assignments; we isolate initial assignments by taking the assignment from this database with the earliest execution date. Using these initial assignments, we fill in assignee organization name as well as assignee address information. Our original source for this data is a set of bulk XML files hosted by Google.[12] These records include application numbers, publication numbers, and patent grant numbers.

4. USPTO patent document pre-grant authority files. A very small number (1.025 total) of published USPTO patent applications are "withdrawn" and these observations tend to be inconsistently reported across the various data sets we need. The USPTO patent document pre-grant authority files – an administrative data file hosted on the USPTO website – allows us to exclude all withdrawn applications for consistency.[13] These records are at the publication number level.

5. USPTO PAIR records. Several key variables for our analysis (e.g. examiner name) are available in the USPTO Patent Application Information Retrieval (PAIR) data, which we draw from an administrative data set called the Patent Examination Research Dataset (PatEx).[14] With the exception of 264 published patent

[10] See `http://www.google.com/googlebooks/uspto-patents-applications-biblio.html`.

[11] See `http://www.google.com/googlebooks/uspto-patents-grants-text.html`.

[12] See `https://www.google.com/googlebooks/uspto-patents-assignments.html`.

[13] See `http://www.uspto.gov/patents/process/search/authority/`.

[14] See `http://www.uspto.gov/learning-and-resources/electronic-data-products/patent-examination-research-dataset-public` for the underlying PAIR data, see: `http://portal.uspto.gov/pair/PublicPair`.

applications, this data is available for our full sample of published USPTO patent applications. These records are at the application number level.

6. Thomson Innovation database. All of the databases listed above record information from the USPTO. One measure of patent value that cannot be constructed based on the USPTO records alone is a measure of patent family size, as developed in Jonathan Putnam's dissertation (Putnam, 1996). Generally stated, a patent "family" is defined as a set of patent applications filed with different patenting authorities (e.g. US, Europe, Japan) that refer to the same invention. The key idea is that if there is a per-country cost of filing for a patent, firms will be more likely to file a patent application in multiple countries if they perceive the patent to have higher private value. Past work - starting with Putnam (1996) - has documented evidence that patent family size is correlated with other measures of patent value. The Thomson Reuters Innovation database collects non-US patent records, and hence allows for the construction of such a family size measure.[15] We purchased a subscription to the Thomson Innovation database, and exported data from the web interface on all available variables for all published USPTO patent applications. To construct our family size measure, we take the DWPI family variable available in the Thomson Innovation database – which lists family members – and separate the country code from the beginning of each number (e.g. "US" in "US20010003111"), and then count the number of unique country codes in the family. These records are at the publication number level.

7. Hall, Jaffe and Trajtenberg (2001) NBER data. Hall, Jaffe and Trajtenberg (2001) constructed a match between US patents granted between January 1963 and December 1999 with the Compustat data. As part of that work, the authors constructed technology categories to describe the broad content area of different patents, based on aggregations of the patent technology class and subclass variables.[16] We match on these technology categories, and hand-fill the small number of cases in which classes or subclasses appear in our data but not in the crosswalk constructed by Hall and co-authors. These records are at the patent class/subclass level.

A.2 Construction of patent application sample

We restrict the sample to USPTO patent applications filed on or after 29 November 2000 (the date when "rejected" applications started to be published), and ends with applications published on 31 December 2013. We impose a few additional sample restrictions:

- We exclude a very small number of "withdrawn" patent applications (1,025 total) given that these observations tend to be inconsistently reported across datasets.

- Six publication numbers are listed in the USPTO patent document pre-grant authority files but are not available in any of our other datasets;[17] we exclude these observations from our sample.

- Four publication numbers are missing from the Thomson Innovation database.[18] We include these observations in the sample, but they are missing data for all variables drawn from the Thomson Innovation data.

- Based on the kind code variable listed on the USPTO published patent applications,[19] we exclude a small number of patent applications that are corrections of previously published applications: corrections of published utility/plant patent applications (kind codes A9/P9; 3,156 total), and second or subsequent publications of the same patent application (kind codes A2/P4; 1,182 total). These kind codes more generally allow

[15] See `http://thomsonreuters.com/thomson-innovation/`.

[16] See `http://www.nber.org/patents/`.

[17] Specifically, these publication numbers are: US20010003111; US20020011585; US20020054271; US20020084413; US20020084764; and US20020103782.

[18] Specifically, these publication numbers are: US20010020331; US20010020666; US20010021099; and US20010021102.

[19] For a summary of USPTO kind codes, see: `http://www.uspto.gov/patents/process/search/authority/kindcode.jsp`.

us to confirm that our sample does not include various types of documents: statutory invention registration documents (kind code H1), reexamination certificates (kind codes Bn/Cn/Fn for n=1-9), post grant review certificates (kind codes Jn for n=1-9), inter parties review certificates (kind codes Kn for n=1-9), or derivation certificates (kind codes On for n=1-9). Our final sample includes only utility patent applications (kind code A1; 3,597,787 total) and plant patent applications (kind code P1; 4,196 total).

Finally, there are two data inconsistencies that we have resolved as follows:

- Seven observations appear to be missing from Google's XML files of the published patent applications.[20] We were able to hand-code the required variables for these observations based on the published patent applications posted at `http://patft.uspto/gov` for all but three of these observations (specifically, publication numbers US20020020603; US20020022313; US20020085735). For those three observations, we hand-coded the required variables based on the information available at `http://portal.uspto.gov/pair/-PublicPair/`; for these, we assumed that the appropriate correspondent addresses were those listed in the "Address and Attorney/Agent," under Correspondence Address.

- When there are multiple published patent applications for a given application number, in order to retain all assignee, correspondent, and inventor information we have appended all assignees/correspondents/inventors into one application number. For example, application number 09878111 has two publication numbers: 20010055990 and 20020058546. Publication number 20010055990 and publication number 20020058546 list "Acres Gaming Incorporated" as the assignee organization. When collapsing this information to the application level, we append these two values into "Acres Gaming Incorporated | Acres Gaming Incorporated" as the assignee organization name for application number 09878111. This issue only arises for three application numbers.[21]

A.3 Description of US Treasury tax files

All variables are constructed from annual business tax returns over the years 1997-2014: C-Corporations (Form 1120), S-Corporations (Form 1120-S), and Partnerships (Form 1065).

Variable Definitions[22]

- Revenue
  - Line 1c of Form 1120 for C-Corporations, Form 1120S for S-Corporations, and Form 1065 for partnerships.
- Total Income
  - For C-Corporations, line 11 on Form 1120. Note that this subtracts COGS from revenues and includes income from a variety of sources (e.g., dividends, royalties, capital gains, etc). For S-Corporations, line 6 on Form 1120S. For partnerships, line 8 on Form 1065.
- Labor Compensation

[20] Specifically, the missing publication numbers are: US20010020331; US20010020666; US20010021099; US20010021102; US20020020603; US20020022313; US20020085735.

[21] Specifically, application numbers 09726661, 09728294, and 09878111. For one of these application numbers (09726661), a patent was granted after the first published application but before the second application was published; for that observation, we code our variable on whether the patent was granted prior to publication as 0.

[22] This section follows the definitions of Yagan (2015) closely. He does not consider partnerships, but the definitions for partnerships are often analogous to those for C-Corporations and S-Corporations.

  - Sum of salaries and wages paid to non-officer employees, employee benefit programs (such as health insurance), and contributions to pension and profit-sharing plans.
  - For C-Corporations, sum of lines 13, 24, 25, and Schedule A line 3 on Form 1120. For S-Corporations, sum of lines 8,17,18, and Schedule A line 3 for Form 1120S. For partnerships, sum of lines 9, 18, 19 on Form 1065 and Form 1125-A line 3. These lines are salaries and wages (other than to partners), retirement plans, employment benefit programs, and the cost of labor on the costs of good sold Form 1125-A, respectively.

- Compensation to Officers

  - For C-Corporations, line 12 on Form 1120. For S-Corporations, line 7 Form 1120. For Partnerships, Line 10 on Form 1065.

- Value Added

  - Gross receipts minus the difference between cost of goods sold and cost of labor.
  - For C-Corporations, line 3 on Form 1120. For S-corporations, line 3 on Form 1120S. For partnerships, line 1c minus line 2 on Form 1065 plus line 3 on Form 1125-A.

- Profits

  - Yagan (2015) defines operating profits as revenues less Costs of Goods Sold and deductions where deductions are total deductions other than compensation to officers, interest expenses, depreciation, and domestic production activities deduction. We do not add back compensation to officers.
  - For C-Corporations, the sum of lines 1c, 18,20 and 25, less the sum of 2 and 27 on Form 1120. We set profits to missing if 1c, 18,20,2, and 27 are all equal to zero.
  - For S-Corporations, operating profits are the sum of lines 1c, 13, and 14 less the sum of 2 and 20 on Form 1120S.
  - For partnerships, operating profits are the sum of lines 1c, 15, and 16c less the sum of 2 and 21 on Form 1065.

- Tangible capital assets (“capital”)

  - Book value of tangible capital assets (excludes e.g., goodwill) that the corporation owns at the end of the year. Capital is net of accumulated book depreciation.
  - Column (d) line 10b on Schedule L of Form 1120 for C-Corporations, Form 1120S for S-Corporations, and Form 1065 for partnerships.

- Assets

  - Book assets includes financial assets like cash, inventories, tangible assets, and intangible assets like goodwill.
  - Column (d) line 15 of Schedule L of Form 1120 for C-Corporations and Form 1120S for S-Corporations. For partnerships, its column (d) line 14 of Schedule L.

- Firm Age

  - Uses the date incorporated field in Box C of Form 1120 for C-Corporations, in Box E of Form 1120S for S-Corporations and Form 1065 for partnerships.

- State
  - Uses the state from the business's filing address.
- Entity Type
  - Indicator based on tax-form filing type.
- Industry
  - NAICS codes are line 21 on Schedule K of Form 1120 for C-Corporations. Line 2a Schedule B of Form 1120S for S-Corporations. For partnerships, Box A of Form 1065.
- Active Firm
  - An active firms has non-zero and non-missing total income and non-missing total deductions.

A.4 Description of merge between patent applications data and US Treasury tax files

Our analysis relies on a new merge between published patent applications submitted to the US Patent and Trademark Office (USPTO) and US Treasury tax files. This data appendix describes the details of this merge, which relies on a fuzzy matching algorithm to link USPTO assignee names with US Treasury firm names.

A.4.1 Creating standardized names within the patent data

Published patent applications list an assignee name, which reflects ownership of the patent application. Due to e.g., spelling differences, multiple assignee names in the USPTO published patent applications data can correspond to a single firm. For example, "ALCATEL-LUCENT U.S.A., INC..", "ALCATEL-LUCENT USA, INCORPORATED," and "ALCATEL-LUCENT USA INC" are all assigned the standardized name "alcatel lucent usa corp".

We employed a name standardization routine as follows. Starting with names in unicode format, we transform the text into Roman alphabet analogs using the "unidecode" library to map any foreign characters into their applicable English phonemes, and then shift all characters to lowercase.[23] We then standardize common terms that take multiple forms, such as "corp." and "corporation"; these recodings were built on the name standardization routine used by the National Bureau of Economic Research (NBER)'s Patent Data Project, with modifications as we saw opportunities to improve that routine.[24] We additionally eliminate any English articles (such as "a" or "an"), since these appeared to be uninformative in our attempts to uniquely identify entities. We then tokenize standardized names by splitting on natural delimiters (e.g. spaces and commas), after which we remove any non-alphanumeric punctuation. Finally, sequences of single character tokens are merged into a combined token (e.g. "3 m corp" would become "3m corp"). The resultant ordered list of tokens constitutes our standardized entity name. We refer to the USPTO standardized firm name as $SNAME_{USPTO}$.

[23]The unidecode library is available at `https://github/com/iki/unidecode`, and is a direct Python port of the Text::Unidecode Perl module by Sean M. Burke.

[24]The NBER Patent Data Project standardization code is available at `https://sites.google.com/site/patentdataproject/Home/posts/namestandardizationroutinesuploaded`.

A.4.2 Creating standardized names within the US Treasury tax files

In the US Treasury tax files, firms are indexed by their Employer Identification Number (EIN). Each EIN is required to file a tax return for each year that it is in operation. Specifically, we restrict our analysis to firms with valid 1120, 1120s, 1065, or 1065b filings over the years 1997-2013. We apply the same name standardization algorithm to the Treasury firm names that was applied to the USPTO names; We refer to the Treasury standardized firm name as $SNAME_{Treasury}$.

A.4.3 Merging standardized names across the USPTO data and the US Treasury tax files

We then conduct a fuzzy merge of $SNAME_{USPTO}$ to $SNAME_{Treasury}$ using the SoftTFIDF algorithm, which is described below. We use this algorithm to allocate each $SNAME_{USPTO}$ to a single $SNAME_{Treasury}$, provided that match quality lies within a specified tolerance. To choose the tolerance we used a hand coded match of applications to Compustat firms as a validation dataset (see A.4.4). The tolerance (and other tuning parameters) were chosen to minimize the sum of Type I and II error rates associated with matches to Compustat firms.

The resulting firm level dataset has one observation per $SNAME_{Treasury}$ in each year. Each observation lists the number of patent applications submitted in that year (APP_{it}), the average leniency of the examiners assigned to those applications (Z^1_{it}), and various measure of the firm's patent stock (S_{it}).

SoftTFIDF Algorithm Our firm name matching procedure of name $a \in SNAME_{USPTO}$ to name $b \in SNAME_{Treasury}$ works as follows. Among all the words in all the firm names in $SNAME_{USPTO}$ that are close to a given word in a, we pick the word with the highest SoftTFIDF index value, which is a word-frequency weighted measure of similarity among words. We do this for each word in the firm's name. For instance, American Airlines Inc would have three words. We then take a weighted-sum of the index value for each word in the firm name where the weights are smaller for frequent words like "Inc." This weighted sum is the SoftTFIDF value at the level of firm-names (as opposed to words in firm names). We assign a to the firm name b with the highest SoftTFIDF value above a threshold; otherwise, the name a as matchless. This subsection provides details on this procedure and example matches.

SoftTFIDF of Firm-Names A score between groups of words X, Y is given by

$$\text{SoftTFIDF}(X,Y) := \sum_{w \in X} \text{weight}(w,X) \cdot \alpha(w,Y) \tag{5}$$

where $\text{weight}(w,X)$ is a word frequency-based importance weight and $\alpha(w,Y)$ is a word match score that uses a word similarity index. Specifically, the importance weight for the word w in the set of words Z is: $\text{weight}(w,Z) := \frac{\text{tfidf}(w,Z)}{\sqrt{\sum_{w' \in Z} \text{tfidf}(w',Z)^2}}$, where

- $\text{tfidf}(w,Z) := \text{tf}(w,Z) \times \text{idf}(w, \mathscr{Z})$
- $\text{tf}(w,Z) := \frac{n(w,Z)}{\sum_{w' \in Z} n(w',Z)}$
- $\text{idf}(w, \mathscr{Z}) := \log\left(\frac{|\mathscr{Z}|}{|\{Z \in \mathscr{Z} \mid w \in Z\}|}\right)$
- $n(w,Z)$ is the number of occurrences of word w in a set of words Z
- $\mathscr{Z}$ is the set of all words in either $SNAME_{USPTO}$ or $SNAME_{IRS}$

We compute the word match score $\alpha(w,Y)$ for words that are close to those in $SNAME_{USPTO}$. To determine which names are close, we use a Jaro-Winkler distance metric to measure the distance between two strings.

Jaro-Winkler Metric of Distance Between Strings We use this metric since it has been shown to perform better at name-matching tasks (Cohen et al. 2003) than other metrics such as Levenshtein distance, which assigns unit cost to every edit operation (insertion, deletion, or substitution). A key component of the Jaro-Winkler metric is the Jaro metric. The Jaro metric depends on the length of $SNAME_{USPTO}$, the length of $SNAME_{Treasury}$, the number of shared letters, and the number of needed transpositions of shared letters.

Specifically, consider strings $s = s_1 \dots s_K$ and $t = t_1 \dots t_L$ and define $H = \frac{\min\{|s|,|t|\}}{2}$, which is half the smaller of K and L. We say a character s_i is **in common with** t if $\exists j \in [i-H, i+H]$ s.t. $s_i = t_j$. Let s', t' be the ordered sets of **in-common** characters (hence we will re-index). Then define $T_{s',t'} := \frac{1}{2}|\{i \mid s'_i \neq t'_i\}|$. The similarity metic is given by

$$\text{Jaro}(s,t) := \frac{1}{3} \cdot \left(\frac{|s'|}{|s|} + \frac{|t'|}{|t|} + \frac{|s'| - T_{s',t'}}{|s'|} \right).$$

The Jaro-Winkler metric is given by

$$\text{Jaro-Winkler}(s,t) := \text{Jaro}(s,t) + \frac{P'}{10} \cdot (1 - \text{Jaro}(s,t)),$$

where P as the longest common prefix of t and s and then $P' = \max\{P, 4\}$, which is the normalization used in Cohen et al (2003).

Word Match score $\alpha(w,Z)$ We define the word match score as follows:

$$\alpha(w,Z) = \begin{cases} 0 & \text{if } \text{closest}(\theta, w, Z) = \emptyset \\ \max\limits_{w' \in \text{closest}(\theta,w,Z)} \text{weight}(w', Z) \cdot \text{Jaro-Winkler}(w, w') & \text{otherwise} \end{cases}$$

where

$$\text{closest}(\theta, w, Z) := \left\{ v \in Z \mid \forall v' \in Z,\ (\text{Jaro-Winkler}(w,v) \geq \text{Jaro-Winkler}(w,v')) \ \wedge\ \text{Jaro-Winkler}(w,v) > \theta \right\}.$$

In words, we select the word w that is the closest importance-weighted match among words that are close to the word w in Z given closeness threshold θ. The accuracy of this matching procedure, which has also recently been used by Feigenbaum (2015), will likely become clearer after reviewing the following examples and discussion of how we selected the tuning parameters (such as the closeness threshold θ).

Example

USPTO Assignee Name	Compustat Firm Name (best match)	Match Score
amted industries inc.	amsted industries	.9921
advanced life sciences inc.	advanced life sciences holdings inc	.9417
alcatel-lucent canada inc.	alcatel-lucent	.8789
adaptive payments, inc.	adaptive solutions inc	.5792
accu-break technologies, inc.	accu test systems inc	.4826
advantage polymers, llc.	advantage payroll services	.3515
alama group inc	alamo group inc	.1146
alereon, inc.	alere inc	.0072

This table provides a small sample of candidate matches from a USPTO to Compustat match.[25] It shows that similar assignee and firm names receive high match scores. Note that common tokens like "inc" do not have a

[25] We present results using Compustat names instead of Treasury names for disclosure reasons.

substantial impact on the match score. To determine the level of match score that constitutes a match, we tune parameters in a hand matched training sample of Compustat names.

A.4.4 Validation: Compustat-USPTO match

This section describes the hand matching process we used to determine the true mapping of USPTO names to compustat names for a random sample of USPTO names. We describe the hand coding task and how we use the hand coded linkages to select the tuning parameters.

Hand Coding Tasks We hired several workers on Upwork (formerly Odesk) as well as University of Chicago undergraduates to hand match two lists of names. The goal for these workers was to match every name in a source file (a list of 100 randomly selected USPTO names) to a target file of Compustat names or to conclude that there is no matching name in the target file. To increase accuracy, we informed these workers that (1) we had hand-coded several of these names ourselves, (2) every name in the source file would be assigned to multiple workers, and (3) we only accepted reasonably accurate work. We also instructed them to use Google to confirm that matches were true matches. For example, "infinity bio ltd" may seem like a match with "infinity pharmaceuticals inc," but Googling the first reveals that the former is a small Brazilian energy company while the latter is a pharmaceutical company headquartered in the US. If one worker found a match but another did not, we considered the non-empty match to be correct. Overall, we ended up assigning 2,196 assignee names, of which 286 (13%) had matches in the Compustat data.[26]

Using Hand Coding Tasks to Select Tuning Parameters We use these hand-coded linkages to establish the "true mapping" from USPTO names to Compustat names, which enables us to select tuning parameters that minimize the sum of type I and type II errors (relative to these "true linkages.")

We constructed a grid and for each set of parameters on the grid executed a match. We then compared these fuzzy matchings to the "true mapping." Type I errors occur when SoftTFIDF returns a match but either (a) the match is inconsistent with the hand-coded match or (b)the hand coded linkage shows no match at all. Type II errors occur when SoftTFIDF does not return a match but the hand-code process had a match.

The parameters that minimize the sum of these false positive and false negatives are: $\theta = .95$, token type of standardized names (instead of raw names), $P = 0$, and a threshold match score of .9. We remind the reader that the parameter θ governs the threshold similarity for two words to be considered "close." Only "close" words contribute to a match score, hence $\theta = .95$ sets a relatively high cutoff below which two similar words do not increase the match score between two fill names. The prefix $P = 0$ suggests that not boosting scores by a common pre-fix doesn't improve performance, which makes sense given that we block by the first letter already.[27] Finally, the threshold match score of .9 shows that we should only consider names a match if they are very close by our similarity metric. With these parameters our matching procedure had error rates below 6% for both types of errors.

A.4.5 Validation: Individual-inventor match

The Bell et al. (2016) inventor-level merge between patent applications and W2 reports in theory can – via the EINs provided on W2 reports – provide a linkage between patent applications and firms, but ex ante we expect this inventor-based match to measure something conceptually different from a firm-based match. For example, many inventors work at firms that are not the assignee of their patents, in which case we would not expect our

[26]This match rate is sensible: the number of compustat names is roughly 20% the number of assignee names, so this rate is consistent with a reasonable proportion of Compustat firms applying for a patent.

[27]It is computationally infeasible to compare every single entity against every other, so we introduced blocked in order to reduce the sizes of sets being compared against one another. In particular, the target names (either Compustat or Treasury names) are chunked by the first letter of their standardized names and grouped with the source (USPTO) names with the same first letter. Hence we will miss any matches that differ on the first letter.

assignee-based merge to match to the same EIN as Bell et al.'s inventor-based merge. However, the Bell et al. merge nonetheless provides a very valuable benchmark for assessing the quality of our assignee-based merge. Bell et al. generously agreed to share their inventor-based merge with us, and our preliminary results comparing the two linkages provide a second set of evidence supporting the quality of our assignee-based linkage. In the simplest comparison, around 70% of patent applications are associated with the same EIN in the two linkages. The characteristics of this match also look sensible, e.g., the match rates are higher if we limit the sample to patent applications that Bell et al. match to inventors who all work at the same firm. Given that we do not expect a match rate of 100% for the reasons detailed above, we view the results of this second validation exercise as quite promising.

B Appendix Shrinkage Procedure

Ideally, we would like to use each examiner's latent leniency as an instrument. Letting $p_{rm} = P\left(G_j^4 = 1 | EX_j = m, ART_j = r\right)$ and $p_r = P\left(G_j^4 = 1 | ART_j = r\right)$, we can define this leniency as:

$$\ell_{rm}^* \equiv p_{rm} - p_r = P\left(G_j^4(m) = 1 | ART_j = r\right) - \frac{1}{J}\sum_{k=1}^{J} P\left(G_j^4(k) = 1 | ART_j = r\right).$$

Note that if we could observe the $\{\ell_{rm}^*\}$, a least squares fit of G_j^4 on ℓ_{ART_j, EX_j}^* would yield an intercept of zero and a slope of one; this property would hold also within each art stratum. The sample analogue of leniency is given by $\ell_{rm} = \bar{G}_{rm} - \bar{G}_r$. While ℓ_{rm} provides an unbiased estimate of ℓ_{rm}^*, it is a noisy proxy because each examiner handles only $N_{rm} = \sum_{j=1}^{J} 1\left[EX_j = m, ART_j = r\right]$ patent applications in an art stratum. Consequently a regression of ℓ_{rm}^* on ℓ_{rm} in an art unit / year cell r would yield an attenuated regression coefficient $\pi_r < 1$ that depends on the average signal to noise ratio in the examiner means in cell r. The predicted values $\pi_r \ell_{rm}$ from such hypothetical regressions form the best linear predictor (BLP) of ℓ_{rm}^* based upon ℓ_{rm}. The BLP is an optimal predictor in the sense that the art stratum specific shrinkage factors $\{\pi_r\}$ employed maximize the correlation with the variable being predicted.[28] Our goal shall be to estimate these shrinkage factors for use in constructing a more powerful instrument for G_j^4 that shrinks the observed ℓ_{rm} in each art stratum towards zero.

To see how the shrinkage factors can be identified in the absence of data on ℓ_{rm}^*, note that $\pi_r = \frac{Cov(\ell_{rm}^*, \ell_{rm})}{Var(\ell_{rm})} = \frac{Var(\ell_{rm}^*)}{Var(\ell_{rm})}$ where the $Cov(.)$ and $Var(.)$ operators give, respectively, variances and covariances across examiners within a given art-unit / year r. Because ℓ_{rm} is observed, estimating its variance is straightforward. To identify $Var(\ell_{rm}^*)$, note that $Var(\ell_{rm}) = Var(\ell_{rm}^*) + Var(\ell_{rm} - \ell_{rm}^*)$. We can expand this second term as:

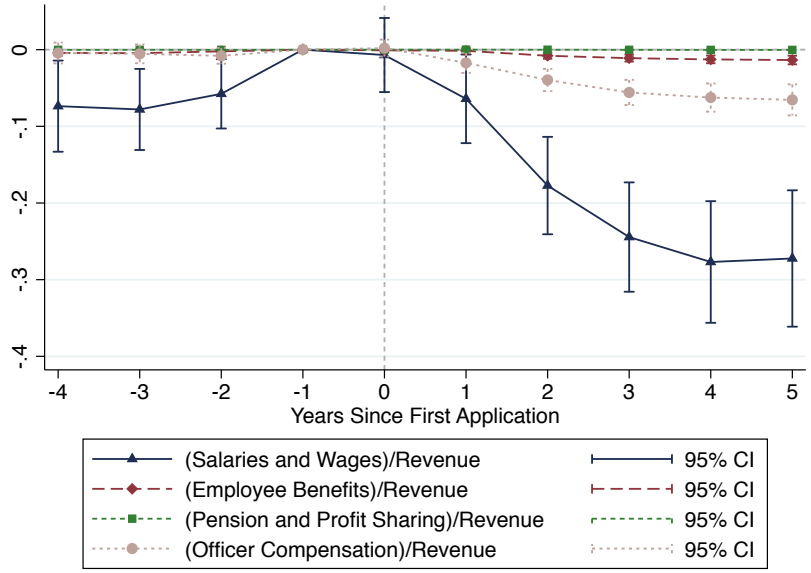
$$\begin{aligned} Var(\ell_{rm} - \ell_{rm}^*) &= E\left[Var(\ell_{rm} - \ell_{rm}^* | N_{rm}, p_{rm}, p_m)\right] + Var\left(E\left[\ell_{rm} - \ell_{rm}^* | N_{rm}, p_{rm}, p_m\right]\right) \\ &= E\left[Var(\ell_{rm} - \ell_{rm}^* | N_{rm}, p_{rm}, p_m)\right]. \end{aligned}$$

To estimate this variance component, we model examiner's decisions as independent Bernoulli trials with success probability p_{rm}. This implies $Var\left(\bar{G}_{rm} | N_{rm}, p_{rm}, p_m\right) = \frac{p_{rm}(1-p_{rm})}{N_{rm}}$, an unbiased estimator for which is $\frac{\hat{p}_{rm}(1-\hat{p}_{rm})}{N_{rm}-1}$. Because there are many examiners in each art stratum we can, as an approximation, ignore sampling uncertainty in $\bar{G}_r$. Hence, $Var(\ell_{rm} - \ell_{rm}^*) \approx E\left[\frac{p_{rm}(1-p_{rm})}{N_{rm}}\right]$, which gives the average amount of noise in the examiner allowance rates in art stratum r. We can estimate this average variance with its unbiased sample analogue $\frac{1}{M_r}\sum_{m=1}^{M_r} \frac{\bar{G}_{rm}\left(1-\bar{G}_{rm}\right)}{N_{rm}-1}$. Therefore, our resulting estimate of the shrinkage factor π_r is $\hat{\pi}_r = 1 - \frac{\frac{1}{M_r}\sum_{m=1}^{M_r} \frac{\bar{G}_{rm}\left(1-\bar{G}_{rm}\right)}{N_{rm}-1}}{\frac{1}{M_r-1}\sum_{m=1}^{M_r} \ell_{rm}^2}$, which is consistent as the number of examiners per art stratum grows large. Our corresponding BLP estimate of ℓ_{rm}^* is given by $\hat{\ell}_{rm} = \hat{\pi}_r \ell_{rm}$.

[28] A standard "random effects" specification would also allow the shrinkage factors to differ across examiners based upon the number of applications N_{rm} handled. In our setting, N_{rm} is likely to be positively correlated with ℓ_{rm}^* – by dispatching cases more quickly a lenient examiner will end up getting assigned more cases. This would imply we should shrink the allowance rate of examiners with more cases towards a different mean, which is an extension we do not consider here.

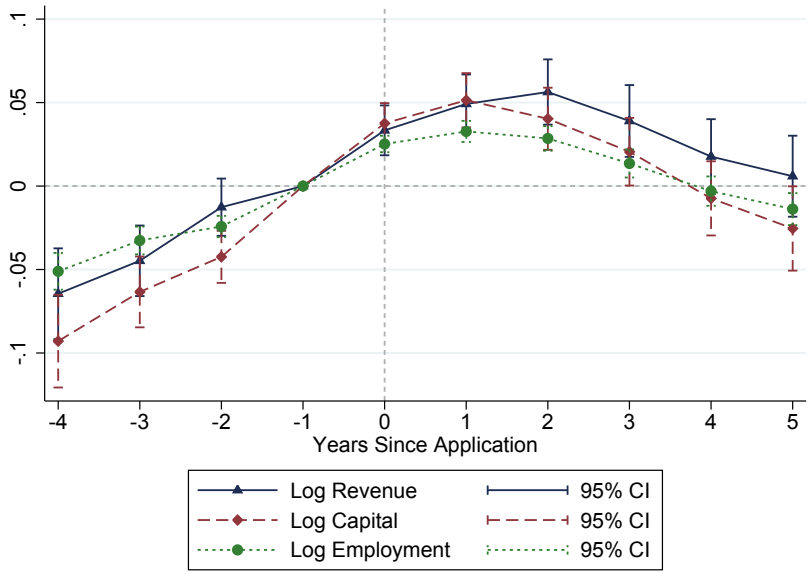
C Appendix Figures

Figure C.1: Event Study of Patent Applications on Worker Compensation by Component



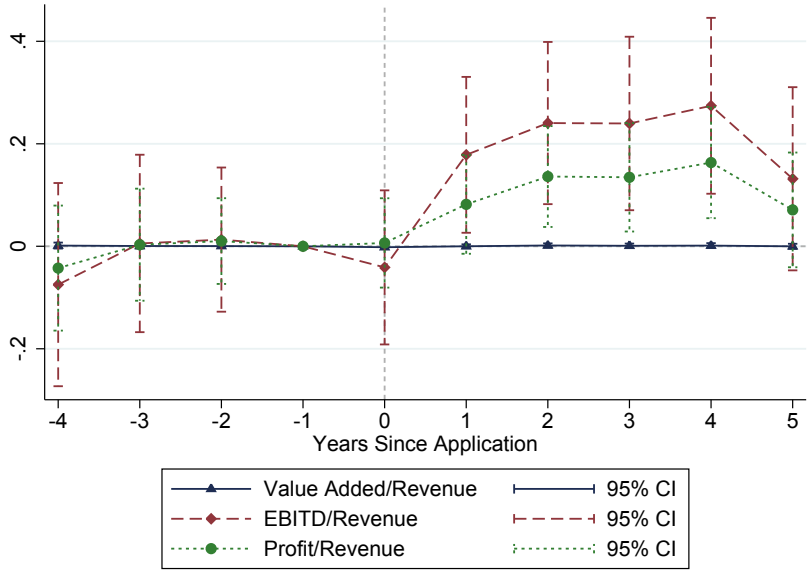
Note: This figure shows estimates of β_k from the event study specification in equation 1. It presents four separate sets of estimates to show how worker outcomes evolve around the firm's first patent application from 2000 to 2010 among infrequently patenting firms, i.e., firms with 10 or fewer patent applications from 2000-2010. Revenue, Compensation, and its four subcomponents – salaries and wages, employee benefits, pension and profit sharing, and officer compensation – are from annual business tax forms F1120, F1120S, and F1065 of active firms (i.e., firms with non-missing and non-zero total income and non-missing total deductions) over the period 1997-2014. Confidence intervals based upon standard errors clustered on firm.

Figure C.2: Event Study of Patent Applications on Firm Size



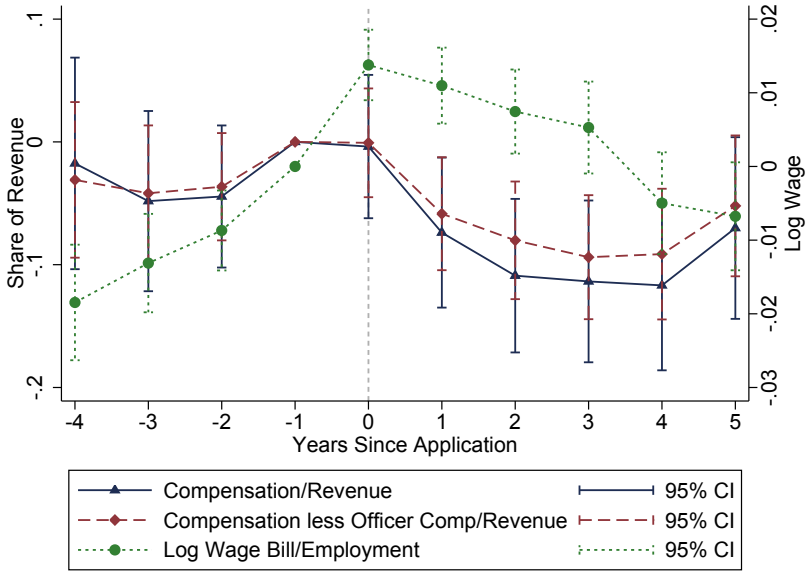
Note: This figure shows estimates of β_k from the event study specification in equation 3. It is similar to Figure 1a but it uses all patent applications among infrequently patenting firms. Specifically, it presents the results of a regression of leads and lags of the number of patent applications on firm size conditional on 2-digit industry-by-year fixed effects. See Figure 1 notes for additional details.

Figure C.3: Event Study of Patent Applications on Firm Performance



Note: This figure shows estimates of β_k from the event study specification in equation 3. It is similar to Figure 2a but it uses all patent applications among infrequently patenting firms. Specifically, it presents the results of a regression of leads and lags of the number of patent applications on firm performance conditional on 2-digit industry-by-year fixed effects. See Figure 2 notes for additional details.

Figure C.4: Event Study of Patent Applications on Worker Outcomes



Note: This figure shows estimates of β_k from the event study specification in equation 3. It is similar to Figure 3a but it uses all patent applications among infrequently patenting firms. Specifically, it presents the results of a regression of leads and lags of the number of patent applications on firm size conditional on 2-digit industry-by-year fixed effects. See Figure 3 notes for additional details.