UC Berkeley
Working Paper Series

Title

Who Profits from Patents? Rent-Sharing at Innovative Firms

Permalink

https://escholarship.org/uc/item/6mr8598q

Authors

Kline, Patrick
Petkova, Neviana
Williams, Heidi
et al.

Publication Date

2017-09-01

Who Profits from Patents? Rent-Sharing at Innovative Firms

Patrick Kline, Neviana Petkova, Heidi Williams, and Owen Zidar

Berkeley
UNIVERSITY OF CALIFORNIA

Who Profits from Patents?

Rent-Sharing at Innovative Firms[*]

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

September 19, 2017

Abstract

This paper analyzes how patent-induced shocks to labor productivity propagate into worker compensation using a new linkage of US patent applications to US business and worker tax records. We infer the causal effects of patent allowances by comparing firms whose patent applications were initially allowed to those whose patent applications were initially rejected. To identify patents that are ex-ante valuable, we extrapolate the excess stock return estimates of Kogan et al. (2017) to the full set of accepted and rejected patent applications based on predetermined firm and patent application characteristics. An initial allowance of an ex-ante valuable patent generates substantial increases in firm productivity and worker compensation. By contrast, initial allowances of lower ex-ante value patents yield no detectable effects on firm outcomes. On average, workers capture 29 cents of every dollar of patent-induced operating surplus. This share is larger for men, employees who are listed as inventors, and firm stayers present since the year of application. Patent allowances lead firms to increase employment, but we find minimal evidence of quality upgrading or selection bias in workforce composition. Surprisingly, entry wages are insensitive to patent decisions, suggesting that the large earnings responses of incumbent workers may reflect performance pay.

1   Introduction

Competitive models of labor markets are predicated on the notion that firms have no power to set wages. However, there is mounting empirical evidence that firms contribute substantially to wage inequality among identically skilled workers (Card, Heining, and Kline 2013; Card, Cardoso, and Kline 2016; Song et al. 2016; Sorkin 2016; Goldschmidt and Schmieder 2017; Helpman et al. 2017; Abowd, McKinney, and Zhao forthcoming; Barth et al. forthcoming). This emerging evidence has renewed interest in theoretical models of firm wage setting that postulate mechanisms through which variation in firm productivity can influence worker pay (see Malcomson 1999; Lentz and Mortensen 2010; Manning 2011 for reviews).

While a sizable empirical literature has documented that fluctuations in firm performance and worker compensation are strongly related (Card et al., forthcoming), these correlations are open to widely varying interpretations. Early studies (e.g., Christofides and Oswald 1992; Blanchflower, Oswald, and Sanfey 1996) estimated industry-level relationships that could simply reflect competitive market dynamics. A second generation of studies (Van Reenen 1996; Hildreth 1998; Abowd, Kramarz, and Margolis 1999) used firm-level data to study how shocks to firm performance translate into worker pay, but was unable to adjust for potential changes in worker composition. More recent work (Guiso, Pistaferri, and Schivardi 2005; Card, Devicienti, and Maida 2014; Card, Cardoso, and Kline 2016; Carlsson, Messina, and Skans 2016; Lamadon 2016) adjusts for composition biases by examining the comovement between changes in firm productivity and the wage growth of incumbent workers. However, observational fluctuations in standard labor productivity measures are likely to reflect a number of factors (e.g., market-wide fluctuations in product demand, changes in non-pecuniary firm amenities, or drift in labor market institutions) that can influence wages without necessarily signaling a violation of price-taking behavior by firms.

In this paper, we investigate the link between firm performance and worker compensation induced by patent allowance decisions. The wage response to patent allowances is of interest for several reasons. First, patents provide firms with well-defined temporary monopoly rights that can yield a prolonged stream of potentially substantial economic rents. Standard models of frictional labor markets (e.g., Pissarides 2000; Hall and Milgrom 2008; Pissarides 2009) suggest that these rents should be shared with workers whenever the employment relationship is (re-)negotiated, yet surprisingly little is known about how broadly such rents are shared in practice. While a handful of studies find that inventor wages increase in response to patent grants (Toivanen and Väänänen 2012; Bell et al. 2016; Aghion et al. 2017), these inventors typically constitute only a small fraction of employment at innovative firms, and the factors governing inventor compensation may not generalize to wider populations. Second, patent allowances fundamentally constitute firm-specific shocks. Hence, using patent allowance decisions as a source of variation in firm performance should effectively filter out market-wide wage responses. Finally, patent allowances are public and verifiable performance measures that innovative firms might plausibly incorporate into pay setting

decisions. While incentive contracts are thought to be common for inventors (Holmström 1989; Aghion and Ti-role 1994; Lerner and Wulf 2007; Manso 2011), there is little evidence on how important performance pay is for workers not directly involved in the inventive process.

Our analysis relies on a new linkage of two datasets: (i) the census of published patent applications submitted to the US Patent and Trademark Office (USPTO) between roughly 2001 and 2011 and (ii) the universe of US Treasury business tax filings and worker earnings histories drawn from W2 and 1099 tax filings. We infer the causal effect of patent allowances by comparing firms whose applications were initially allowed to those whose applications were initially rejected. Within so-called "art units" (technological areas designated by the USPTO), firms with initially allowed and initially rejected applications submitted in the same year are found to exhibit similar levels and trends in outcomes prior to their initial patent decision. We also document that initial patent decisions are difficult to predict based on firm characteristics or geography, corroborating the view that these decisions constitute truly idiosyncratic shocks.

It is well-known that most patents generate little ex-post value to the firm (Pakes 1986; Hall, Jaffe, and Trajten-berg 2001). We build on insights from two recent studies to identify a subsample of valuable patents that induce meaningful shifts in firm outcomes at the time the patents are allowed. First, following the work of Farre-Mensa, Hegde, and Ljungqvist (2017), we restrict our analysis to firms applying for a patent for the first time, for which patent decisions are likely to be more consequential. Second, among this sample of first-time applicants, we build on the analysis of Kogan et al. (2017) who use event studies to estimate the excess stock market return realized on the grant date of US patents assigned to publicly traded firms. Specifically, we develop a methodology for extrapolating Kogan et al.'s patent value estimates to the non-publicly traded firms in our sample, and to firms whose patent applications are never granted. We use characteristics of firms and their patent applications that are fixed at the time of application as the basis for extrapolating patent values, and show that these value estimates are strong predictors of treatment effect heterogeneity in our sample. These value estimates also provide us with an additional validation of our research design: patents with low predicted value are found to have economically small and statistically insignificant effects on firm performance and worker compensation.

Using these data, we then investigate the consequences of obtaining an ex-ante valuable patent allowance for firm performance and worker compensation, and relate our findings to different explanations for the propagation of firm-specific shocks into worker wages. Corroborating recent research based on US Census data (Balasubrama-nian and Sivadasan 2011), we find that firm size and average labor productivity rise rapidly in response to initial allowances of ex-ante valuable patents. The average wage and salary income of workers at these firms rises in tandem with measures of average labor productivity. An allowance of a patent application in the top quintile of ex-ante predicted value raises firm-level operating surplus—defined as the sum of W2 earnings and business earn-ings before interest, taxes, and depreciation—by roughly $12,000 per W2 employee per year, while wage bill per

worker—defined as average W2 earnings at the firm—rises by approximately $3,600 per worker per year.

Patent allowances not only raise average earnings at assignee firms, but also exacerbate within-firm inequality on a variety of margins. The earnings of male employees rise strongly in response to a patent allowance, while the earnings of female employees show little response to patent decisions. We also document that the earnings of "inventors" —defined as employees ever listed as inventors on a patent application as in Bell et al. (2016) —are more responsive to patent allowances than are the earnings of non-inventors. Finally, earnings impacts are heavily concentrated among employees in the top quartile of the within-firm earnings distribution, while bottom quartile earnings are unresponsive to patent decisions. These findings strongly suggest that firms play an important role in generating earnings inequality not only *across* but also *within* workplaces.

Surprisingly, we uncover little evidence that innovative firms change the composition of their workforce in response to patent decisions. While patent allowances lead firms to expand by hiring slightly younger workers, the average prior earnings of both new hires and firm separators is unaffected by patent decisions, suggesting there are no major changes in the skill composition of worker inflows or outflows from the firm on a year-to-year basis. There is some evidence that patent allowances increase the retention rate of workers who were employed at the firm in the year of the patent application. However, these additional retained workers appear similar in terms of application year earnings to workers who would have been present had the firm not been allowed a patent.

Different theoretical frameworks offer divergent predictions about how firm-specific shocks will affect entry wages and the wages of incumbent workers. Empirically, we document that the earnings of workers who were employed by the firm in the year of application respond strongly to patent decisions. Having a valuable patent allowed raises the average earnings of these "firm stayers" by roughly $8,000—or approximately 10%—per year. These gains are widely distributed across firm stayers regardless of their position in the firm's earnings distribution at the time of application. By contrast, we are unable to detect any response of entry wages to patent allowances, which is inconsistent with the predictions of both static wage posting models and traditional bargaining models involving Nash-style surplus splitting at the time of hiring (Pissarides 2000; Hall and Milgrom 2008; Pissarides 2009). While some dynamic wage posting models (e.g., Postel-Vinay and Robin 2002) can generate drops in entry wages in response to a productivity increase, these models predict greater wage *growth* for new hires, a phenomenon for which we also find no evidence.

One candidate explanation for such an "insider/outsider" distinction in earnings impacts is that the wage fluctuations of incumbent workers represent changes in market perceptions of a worker's underlying ability (Gibbons and Murphy 1992; Holmström 1999). Such "career concerns" explanations predict that workers should be able to take some of this earnings advantage with them to new employers. However, we find much smaller and statistically insignificant earnings effects on workers who leave the firm, suggesting that our results are unlikely to be driven by public learning about worker quality. Instead, we argue that the differential response of incumbent workers' wages

is more consistent with the presence of implicit contracts that tie worker pay to performance measures. Models of such contracts (e.g., Bull 1987; MacLeod and Malcomson 1998; Levin 2003) suggest that rewarding *incumbent* workers for performance—either via bonuses of the sort studied by Lemieux, MacLeod, and Parent (2009) or promotions and title changes of the sort documented by Baker, Gibbs, and Holmström (1994)—can serve to mitigate moral hazard and adverse selection problems in employment relationships.

Using patent decisions as an instrument for operating surplus, we then fit a series of "rent-sharing" specifications analogous to cost-price pass-through specifications used to study imperfect competition in product markets (Goldberg and Hellerstein, 2013; Weyl and Fabinger, 2013; Gorodnichenko and Talavera, 2017). We find that worker earnings rise by roughly 29 cents of every dollar of patent allowance-induced operating surplus, with an approximate elasticity of 0.19. Importantly, failing to instrument for operating surplus yields smaller elasticities, closer to those in the recent studies reviewed by Card et al. (forthcoming) that assume statistical innovations to average labor productivity constitute structural productivity shocks. One economic interpretation of this discrepancy is that many observable fluctuations in measured productivity are uninformative about worker effort, leading managers to ignore them when setting pay or making promotion decisions (Holmström 1979, 1989). Consistent with this interpretation, we find that rent-sharing with firm stayers is more pronounced than it is with average workers: stayers capture roughly 60 cents of every dollar of surplus for an approximate elasticity of 0.23.

We conclude with a discussion of the implications of our findings for the literature on rent-sharing and for tax and innovation policy. Our results clearly demonstrate that innovative firms are not passive price takers of the sort traditionally used to interpret trends in wage inequality (e.g., Katz and Murphy 1992; Juhn, Murphy, and Pierce 1993). However, wage responses to firm-specific shocks do not necessarily signal the presence of ex-ante rents or market inefficiency. Indeed, in classic models of efficient incentive contracts (e.g., Mirrlees 1976; Holmström 1979), workers are held to a participation constraint which makes them indifferent between staying at the firm and leaving for an outside market wage. Of course, regardless of how much backloading of wages may have taken place ex-ante, the pronounced wage increases associated with patent allowances are very likely to signal ex-post increases in welfare for incumbent workers. Such "risk-sharing" schemes could provide strong incentives for both inventors and non-inventors to exert effort on activities integral to the innovation and patent application process.

Finally, our findings highlight an important interaction between the tax treatment of firms and workers, particularly relevant to so-called "patent box" proposals that seek to reduce taxes on the profits associated with patents (see Furman (2016) for one discussion). Our results suggest that such subsidies will tend to boost the wage impacts of innovations, thereby increasing the tax revenue raised from individual income taxes. The existence and magnitude of this type of fiscal externality has, to our knowledge, so far been ignored in discussions of patent box policies as well as most other business tax policies.

2 Interpreting Wage Fluctuations

A variety of departures from the textbook competitive labor market model can generate propagation of firm-specific shocks into wages. In this section, we contrast the predictions of alternate models of firm wage-setting, and discuss the ways in which these models interpret wage responses to fluctuations in firm productivity. First we discuss simple wage-posting models in which firms exert market power to extract labor market rents. We then consider an alternative class of models in which firms link pay to performance in order to mitigate moral hazard problems.

2.1 A Wage Posting Model

We begin by developing a variant of the static wage posting model of Card et al. (forthcoming), which attributes wage dispersion across firms to workplace differentiation, to accommodate a continuum of worker types. Despite abstracting from the time-consuming process of search, the model yields a wage-setting rule similar to those found in many search models with multi-lateral bargaining (Pissarides 2000; Cahuc and Wasmer 2001; Acemoglu and Hawkins 2014), as well as in much of the classic literature on union wage bargaining (Brown and Ashenfelter 1986). We use this framework to motivate standard empirical "rent-sharing" specifications and to clarify the endogeneity problems that arise when estimating the transmission of firm-specific shocks to wages. We then discuss the assumptions under which patent allowance decisions can facilitate the identification of economic parameters of interest.

2.1.1 Labor Supply

There is a unit mass of workers of varying quality. The supply to firm j of workers with quality level q is given by:

$$N_j(q) = N(q)(w_j(q) - b_j(q))^{\beta}$$

where $N(q)$ denotes the population density of workers with quality level q and $b_j(q)$ denotes a reservation wage below which no worker of quality q is willing to work at firm j (Manning 2011). Variation in $b_j(q)$ captures both heterogeneity in the value of leisure (or other alternatives) across quality types and any firm-specific amenities. Above the minimal reservation wage, workers supply labor to the firm with a constant elasticity β (i.e., β gives the elasticity of labor supply with respect to the net wage $w_j(q) - b_j(q)$). As described in Card et al. (forthcoming), β can be thought of as an inverse measure of the degree of heterogeneity in worker preferences for firms' unobserved workplace characteristics. There are many such characteristics that could serve to differentiate firms, including location, workplace culture, and the nature of the tasks that the employer requires of workers. Whatever the source

of differentiation, a finite β endows firms with some power to set wages.

The elasticity of supply of type q workers to firm j is given by:

$$\eta_j(q) \equiv \frac{d \ln N_j(q)}{d \ln w_j(q)} = \beta \frac{w_j(q)}{w_j(q) - b_j(q)}. \tag{1}$$

This elasticity is very large when the wage is near the reservation value $b_j(q)$, but grows smaller at higher wage levels. As $\beta \to \infty$, labor supply becomes perfectly elastic around the (firm-specific) reservation wage as in competitive models with compensating differentials (Rosen 1986).

2.1.2 Production and Wage Rule

Firms produce output Q_j from workers according to the production function:

$$Q_j = T_j \bar{q}_j N_j$$

where T_j denotes firm j's total factor productivity (TFP), $\bar{q}_j$ denotes the average quality of the firm's employees, and N_j denotes the total number of workers employed at the firm. This efficiency units formulation assumes workers of quality levels are perfect substitutes in production, which simplifies estimation of average labor productivity. While workers of different quality levels are likely to be imperfect substitutes in practice, the increasingly stark skill segregation of workers across US firms (Song et al. 2016) likely makes this a reasonable working model for studying firm responses to productivity shocks.

The firm chooses a quality-dependent wage schedule $w(q)$ to minimize costs given knowledge of the labor supply schedule $N_j(q)$ and the production function. Formally, the firm's problem is to:

$$\min_{w(q)} \int w(q) N_j(q) dq \text{ s.t. } Q_j = \bar{Q}.$$

The first-order condition for an interior solution to this problem is a quality dependent pricing rule:

$$w_j(q) = \frac{\eta_j(q)}{1 + \eta_j(q)} T_j q \mu_j, \tag{2}$$

where $\mu_j = \mu(Q_j)$ is a Lagrange multiplier capturing the marginal cost of production. Equation (2) implies the usual monopsony markdown rule of wages below marginal revenue product $T_j q \mu_j$ (Robinson 1933; Manning 2011). This condition will hold for any quality level with positive employment at the firm.

An optimizing firm will equate the marginal cost of production to marginal revenue. We assume that all firms

6

face product demand curves with common price elasticity ε, in which case marginal revenue can be written

$$\mu_j = P_j \left(1 - \frac{1}{\varepsilon}\right),$$

where P_j is firm j's product price, which exhibits a fixed markup of $\frac{\varepsilon}{\varepsilon-1}$ over its marginal cost μ_j.

Plugging equation (1) into equation (2) and simplifying yields a linear wage rule:

$$w_j(q) = (1 - \theta) b_j(q) + \theta T_j q \mu_j \tag{3}$$

where $\theta = \frac{\beta}{1+\beta}$. Workers of each type are paid a θ-weighted average of their marginal productivity $T_j q \mu_j$ and the reservation wage $b_j(q)$. Rewriting $\theta = \frac{w_j(q) - b_j(q)}{T_j q \mu_j - b_j(q)}$ illustrates the link to models with Nash wage bargaining in which θ gives the fraction of marginal match surplus $T_j q \mu_j - b_j(q)$ paid out in wage premia $w_j(q) - b_j(q)$.[1] As the labor supply parameter β increases, θ rises and workers capture more of the surplus.

The parameter θ has a clear causal interpretation: a dollar increase in marginal productivity yields a θ cent pay increase. Note that marginal products vary across firms here because workplaces are imperfect substitutes in the eyes of workers, whereas in the competitive model, firms take wages as given and scale is adjusted until marginal products are equalized.

Averaging equation (3) across workers at the firm level yields:

$$
\begin{aligned}
\bar{w}_j &= (1 - \theta) \bar{b}_j + \theta T_j \bar{q}_j \mu_j \\
&= (1 - \theta) \bar{b}_j + \theta \left(1 - \frac{1}{\varepsilon}\right) \frac{P_j Q_j}{N_j} \\
&= (1 - \theta) \bar{b}_j + \pi \bar{S}_j
\end{aligned}
\tag{4}
$$

where $\bar{b}_j = N_j^{-1} \int b_j(q) N_j(q) \, dq$ denotes the average reservation value among firm employees. The last line of this expression is a standard empirical rent-sharing specification relating average wages at the firm to a measure of average labor productivity $\bar{S}_j = \frac{P_j Q_j}{N_j}$, which we refer to as *gross surplus* per worker. The parameter $\pi = \theta \left(1 - \frac{1}{\varepsilon}\right)$ governs pass-through of gross surplus to wages and can be thought of as the labor market analog of cost-price pass-through coefficients often used to study imperfect competition in product markets (Goldberg and Hellerstein 2013; Weyl and Fabinger 2013; Gorodnichenko and Talavera 2017). The term $\left(1 - \frac{1}{\varepsilon}\right)$ is an adjustment factor that converts average labor productivity to marginal labor productivity. While π is our primary parameter of interest, we also explore calibrations of ε and consider the implied values of the structural rent-sharing coefficient θ.

---

[1]Stole and Zwiebel (1996) propose a multilateral bargaining framework where workers and firms also bargain over infra-marginal products. This bargaining concept is embedded in a search and matching framework by Acemoglu and Hawkins (2014). Given our assumption of a constant product demand elasticity, the wage rule that results from that approach is analogous to equation (3) with the modification that the weights on the reservation wage and marginal revenue product need not sum to 1.

2.1.3 Threats to Identification

The primary empirical challenge to estimating the pass-through coefficient π in equation (4) is that we cannot observe the average reservation value $\bar{b}_j$ of firm employees, which could plausibly be correlated with the average labor productivity $\bar{S}_j$ for a number of reasons. For example, firms could engage in skill upgrading in response to an increase in labor productivity. We show in Appendix B that this possibility happens when reservation wages are convex in labor quality. Skill upgrading leads $\bar{b}_j$ and $\bar{S}_j$ to be positively correlated, which would bias OLS estimates of π upwards.

A related concern is that shocks to firm productivity may contain a market-wide component. If all firms in a market become more productive, reservation wages will rise. This possibility would lead to a misattribution of market-level wage adjustments to rent-sharing and a corresponding upward bias in OLS estimates of π.

A different class of potential biases arises from unobserved shocks to the amenity value of a firm. Suppose the work environment at a firm improves and leads to a decrease in $\bar{b}_j$. This improvement will lead, ceteris paribus, to an increase in firm scale, which will tend to depress average labor productivity through drops in the product price P_j. Consequently, such shocks will induce a positive covariance between $\bar{b}_j$ and $\bar{S}_j$ and hence lead to an overstatement of the degree of rent-sharing. It seems reasonable to expect, however, that unobserved amenity shocks could also exert a direct effect on productivity. For example, a recent empirical literature finds that variation in management practices affect both worker morale and productivity (Bloom and Van Reenen 2007; Bender et al. forthcoming). A new manager who motivates workers could plausibly raise total factor productivity T_j while lowering reservation wages. This possibility would lead to an under-estimate of rent-sharing as the productivity shock is accompanied by an unobserved amenity shock.

2.1.4 Instrumenting with Patent Decisions

To circumvent these endogeneity problems, we use the initial decision of the US Patent and Trademark Office (USPTO) on a firm's first patent application as an instrument for average labor productivity.[2] A priori, patents could influence average labor productivity through at least two channels, both of which provide valid identifying variation. First, a patent grant could allow a firm to raise its product price P_j by creating a barrier to competition by rival firms.[3] Second, a patent grant could raise a firm's TFP T_j by making it profitable for the firm to implement the patented technology.

---

[2]Van Reenen (1996) also investigated patents as a source of variation, but found them to be a relatively weak predictor of firm profits in his sample of firms (see his footnote 11). This finding is in keeping with the notion that most patents generate little ex-post value to the firm (Pakes 1986).

[3]Perhaps the classic example is patents on branded small molecule pharmaceuticals. In the absence of patents, many branded pharmaceuticals would experience near-immediate entry of generic versions which compete with branded pharmaceuticals at close to marginal cost prices.

We document below that within observable strata, the USPTO's initial decision on a given patent application is unrelated to trends in firm performance, implying that initial patent decisions are as good as randomly assigned with respect to counterfactual *changes* in firm outcomes. Consistent with this evidence, we also document below that it is hard to predict initial decisions using firm characteristics in the year of application.

We further assume that patent decisions do not directly affect the amenity value $b_j(q)$ of the firm. This assumption might be violated for two (unmodeled) reasons. One is that patent allowances might lead the firm to demand more hours from workers, in which case $b_j(q)$ would rise. However, we would expect this to be a short-run phenomenon that dissipates as the firm expands towards its new target size, and we find no evidence of such wage dynamics in the data. A different sort of violation would occur if patents shift expectations about firm growth and therefore about the future earnings growth of workers. This sort of mechanism arises in dynamic wage posting models with offer matching (Postel-Vinay and Robin 2002) and would imply that $b_j(q)$ falls in response to an allowance. However, such a violation would also imply that initial allowances should raise the wage growth of new hires, an assertion for which we find no empirical support.

The assumption that patent decisions do not affect $b_j(q)$, in conjunction with random assignment, implies that our instrument filters out management shocks or drifts in the amenity value of the firm. We also document below that initial allowance decisions are not geographically correlated, which casts doubt on the potential for market-level wage movements to bias our instrumental variables estimates.

However, at least two sources of potential biases remain. First, granting a patent to a firm could harm rival firms. If those rival firms also employ similar workers, this competition could lead to a downward bias in estimates of π by lowering $\bar{b}_j$. While we cannot formally rule out this possibility, it seems unlikely that product market rivals employ a large fraction of the workers who are considering employment at any given firm, suggesting that the magnitude of this type of bias is small.

Second, firms may respond to patent decisions by changing the composition of their workforce. By leveraging the panel structure of our data, we can examine whether firms change their composition of new hires (or separations) in response to patent allowances. We can also composition adjust our rent-sharing estimates by analyzing the wage *growth* of firm stayers, defined as workers who have been employed at the firm since the time of the patent application.

2.2 Performance Pay Models

The monopsony model explains wage fluctuations as attempts to adjust firm scale. A different explanation for wage fluctuations is that firms link pay to performance to address moral hazard problems. In the simplest example, a firm might incentivize workers to innovate by offering a "bonus" if a patent is granted. This bonus component of

pay is a reward to incumbent employees for work already done and therefore might not result in a change in firm size (or entry wages) if the environment is stationary.

While formal incentive contracts are more likely to be found among inventors and executives, a well-developed theoretical literature shows more generally that when employment relationships are long-lived, "self-enforcing" implicit contracts that link pay and promotion decisions to observable performance measures may emerge (Bull 1987; MacLeod and Malcomson 1998; Levin 2003). Earnings fluctuations driven by such contracts need not signal the presence of ex-ante rents or inefficiency. Indeed, in standard principal-agent models (Mirrlees 1976; Holmström 1979), contracts are designed to ensure that the agent is indifferent about taking the job and behavior is (constrained) efficient. Nevertheless, the optimal contract generally does involve some ex-post sharing of rents that generates risk and contributes to earnings inequality. Empirically, our key attempt to distinguish between this class of theoretical models and the monopsony model will involve contrasting the response of incumbent workers' wages with the response of entry wages. Notably, a failure of entry wages to respond to a valuable patent allowance would represent a departure not only from the static wage posting model described above but also from standard bargaining models, which typically predict that workers will negotiate a better wage at the time of hiring when the firm is doing well (Pissarides 2000; Hall and Milgrom 2008; Pissarides 2009).

Performance pay models also provide an interesting explanation for discrepancies between IV and OLS estimates of the pass-through parameter π. An optimizing firm should only tie pay to signals that are informative about worker effort (Holmström 1979). If, as seems likely, many fluctuations in firm productivity are uninformative about worker effort, we expect OLS estimates of π to be smaller than estimates where an informative signal is used as an instrument for firm performance. Evidence from Kogan et al. (2017) establishes that patent decisions are highly visible to investors. It seems plausible that they are also potentially informative about effort, leading to the presumption that instrumenting average labor productivity with patent decisions should raise the estimated value of π. By contrast, our above discussion of the monopsony model suggested it was more likely that OLS would over-estimate π. Hence, a comparison of our IV and OLS estimates will provide a second source of evidence for distinguishing between performance pay mechanisms and monopsony power.

3 Data and Descriptive Statistics

To conduct our empirical analysis, we construct a novel linkage of several administrative databases, which provides us with panel data on the patent filings, patent allowance decisions, and outcomes of US firms and workers.

3.1 USPTO Patent Applications

We begin with public-use administrative data on the universe of patent applications submitted to the US Patent and Trademark Office (USPTO) since late 2000.[4] We link these published US patent applications with several USPTO administrative datasets. Because published patent applications are not required to list the assignee (owner) of the patent, approximately 50% of published patent applications were originally missing assignee names. We worked with the USPTO to gain access to a separate public-use administrative data file that allows us to fill in assignee names for most of these names. The public-use USPTO PAIR (Patent Application Information Retrieval) administrative data records the full correspondence between the applicant and the USPTO, allowing us to infer the timing and content of the USPTO's initial decision on each patent application as well as other measures of USPTO and applicant behavior. Details on these and the other patent-related data files that we use are included in Appendix A.

Panel A of Table 1 describes the construction of our patent application sample. Our full sample consists of the roughly 3.6 million USPTO patent applications filed on or after 29 November 2000 that were published by 31 December 2013; we restrict attention to applications filed on or before 31 December 2010 in order to limit the impact of censoring. We drop around 400,000 applications that are missing assignee names and therefore cannot be matched to business tax records. We also limit our sample to standard (so-called "utility") patents.[5]

To focus on a subset of firms for which patent allowances are most likely to induce a meaningful shift in firm outcomes, we make several restrictions that aim to limit our sample to first-time patent applicants. First, we drop so-called "child" applications that are derived from previous patent applications. Second, we retain the earliest published patent application observed for each assignee in our sample.[6] Finally, we exclude assignees which we observe to have had patent grants prior to the start of our published patent application sample. Ideally, we would exclude assignees that had patent *applications* (not just patent grants) prior to the start of our published patent application sample, but unsuccessful patent applications filed before 29 November 2000 are not publicly available. These restrictions leave a sample of around 96,000 patent applications, which we then attempt to match to our US Treasury business tax files.

---

[4]The start date of our sample is determined by the American Inventors Protection Act of 1999, which required publication of nearly the full set of US patent applications filed on or after 29 November 2000. We say "nearly" because our sample misses patent applications that opt out of publication; Graham and Hegde (2014) use internal USPTO records to estimate that around eight percent of USPTO applications opt out of publication.

[5]Utility patents, also known as "patents for invention," comprise approximately 90% of USPTO-issued patent documents in recent years; see https://www.uspto.gov/web/offices/ac/ido/oeip/taf/patdesc.htm for details.

[6]Because USPTO procedure assigns application numbers sequentially, we break ties in the cases in which a given assignee submits multiple applications on the same day by taking the smallest application number.

3.2 Treasury Tax Files

We link US Treasury business tax filings with worker-level filings. Annual business tax returns record firm outcomes from Form 1120 (C-Corporations), 1120S (S-Corporations), and 1065 (Partnership) forms, and cover the years 1997-2014. The key variables that we draw from the business tax return filings are revenue, value added, EBITD (earnings before interest, taxes, and deductions), and labor compensation; each of these is defined in more detail in Appendix A.

We link these business tax returns to worker-level W2 and 1099 filings in order to measure employment and compensation for employees (e.g., wage bill) and independent contractors, respectively, at the firm-year level. The relevant variables are defined in more detail in Appendix A. We winsorize all monetary values in the tax files from above and below at the five percent level, which is standard when working with the population of US Treasury business tax files (see, for example, DeBacker et al. 2016). Since our analysis focuses on per-worker outcomes, we winsorize outcomes on a per-worker basis.

To distinguish employment and compensation for inventors and non-inventors, we use Bell et al.'s (2016) merge of inventors listed in patent applications to W2 filings. Inventors are defined as individuals ever appearing in the Bell et al. (2016) patent application-W2 linkage, rather than individuals listed as inventors on the specific patent application relevant to a given firm.

3.3 Linkage Procedure

We build on the name standardization routine used by the National Bureau of Economic Research (NBER)'s Patent Data Project (`https://sites.google.com/site/patentdataproject/`) to implement a novel firm name-based merge of patent assignees to firm names in the US Treasury business tax files. Specifically, we standardize the firm names in both the patent data and (separately) the US Treasury business tax files 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 conduct a fuzzy merge of standardized assignee names to standardized firm names in the business tax files using the SoftTFIDF algorithm based on a Jaro-Winkler distance measure. This merge is described in more detail in Appendix A.

To assess the quality of our merge, we conducted two quality checks: first, we validate against a hand-coded sample; and second, we validate against the inventor-based linkage of Bell et al. (2016). As described in Appendix A, the results of these validation exercises suggest that our merge is of relatively high quality, with type I and II error rates on the order of five percent.

Panel B of Table 1 describes our linkage between the USPTO patent applications data and the US Treasury business tax files. Of the around 96,000 patent applications we attempt to match to the US Treasury business tax

files, we match around 40,000 patent applications. Given that the USPTO estimates that in 2015 approximately 49.6% of USPTO patent grants were filed by US-based assignees, our implied match rate to US-tax-paying entities is on the order of 83%.[7] These 40,000 patent applications are matched to around 40,000 standardized firm names in the US Treasury business tax files, which corresponds to 82,000 firms (employer identification numbers, or EINs).

We build the analysis sample from these 82,000 EINs in four steps. Our goal here is to construct a unique and well-defined match between patent applications and firms in a subset of firms for which patent allowances are most likely to induce a meaningful shift in firm outcomes. First, we attempt to restrict our post-merge tax analysis sample to first-time patent applicants by retaining the earliest-published patent application observed for each EIN, and by excluding EINs which we observe to have had patent grants prior to the start of our published patent application sample. Second, in cases in which there are multiple EINs for a standardized name in the tax files, we keep the EIN with largest revenue in the year that the patent application was filed. Third, we restrict attention to "active" firms, defined as EINs that have a positive number of employees in the year of application and non-zero, non-missing total income or total deductions in the year the patent application was filed and in the three previous years. This restriction allows us to investigate pre-trends in our outcome variables among economically relevant firms. Fourth, we limit attention to EINs with less than 100 million in revenue in 2014 USD in the year of patent application. This step, which eliminates firms in the top centile of the firm size distribution, allows us to avoid complexities related to the largest multinational companies and focus on firms for whom patent allowance decisions are more likely to be consequential.[8] These restrictions leave us with a sample of 9,735 patent applications, each uniquely matched to one EIN in the US Treasury business tax files. It is worth noting that focusing on such a small subset of firms is common in analyses such as ours. For example, Kogan et al. (2017) start with data on 7.8 million granted patents, which they winnow down to a final sample of 5,801 firms with at least one patent.

3.4 Measuring Surplus

As described in Card et al. (forthcoming), empirical rent-sharing estimates are often sensitive to a number of measurement issues, the most prominent of which is the choice of rent measure. In keeping with equation (4), we rely on a *gross surplus* measure meant to capture the sum of a firm's profits and its wage bill (which we define as total W2 earnings in that firm-year). As discussed in Section 2, this concept—which we also refer to as "operating surplus"—differs from match surplus due to the absence of data on workers' reservation wages. We measure a firm's operating surplus as the sum of its wage bill and its earnings before interest, tax, and depreciation (EBITD). Though firms sometimes report negative EBITD, operating surplus is typically positive and provides a plausible

---

[7]These USPTO estimates, which are based on the reported location of patent assignees, are available here: https://www.uspto.gov/web/offices/ac/ido/oeip/taf/own_cst_utl.htm.

[8]Statistics for firm size distribution are from Smith et al. (2017). Specifically, in the full population of C-corporations, S-corporations, and Partnerships with positive sales and positive W2 wage bills, $100 million in revenue in 2014 USD falls in the top one percent of firms.

upper bound on the flow of resources that could potentially be captured by workers.

3.5 Summary Statistics

Table 2 tabulates summary statistics on our firm and worker outcomes in each of two samples: our analysis sample of matched patent applications/firms (N=9,735), and our sub-sample of matched patent applications/firms for which the patent applications are in the top quintile of predicted value (N=1,947), which will be defined in the next section. All summary statistics are as of the year the patent application was filed.

Panel A documents summary statistics on firm-level outcomes. In our analysis sample, the median firm generated around three million dollars in revenue, employed 17 workers, and reported roughly $7,000 in EBITD per worker. Approximately 8% of patent applications are initially allowed. Panel B documents summary statistics on worker-level outcomes. The median firm in our analysis sample paid $48,000 in annual earnings per W2 employee, employed a workforce that is approximately 75% male, and issued 2.5% of its W2s to individuals listed as inventors on at least one patent application. Contract work turns out to be relatively uncommon in this sample, with 1099s constituting only about 10% of the sum of W2 and 1099 employment for the median firm.

4 Institutional Context: Initial Patent Decisions

The US Patent and Trademark Office (USPTO) is responsible for determining which—if any—inventions claimed in patent applications should be granted a patent. Patentable inventions must be patent-eligible (35 U.S.C. §101), novel (35 U.S.C. §102), non-obvious (35 U.S.C. §103), useful (35 U.S.C. §101), and the text of the application must satisfy the disclosure requirement (35 U.S.C. §112). When patent applications are submitted to the USPTO, they are routed to a central office which directs the application to an appropriate "art unit" that specializes in the technological area of that application. For example, art unit 1671 reviews applications related to the chemistry of carbon compounds, whereas art unit 3744 reviews applications related to refrigeration. The manager of the relevant art unit then assigns the application to a patent examiner for review. If the examiner issues an initial allowance, the inventor can be granted a patent. If the examiner issues an initial rejection, the applicant has the opportunity to "revise and resubmit" the application, and the applicant and examiner may engage in many subsequent rounds of revision (see Williams 2017 for more details).

Our empirical strategy focuses on contrasting firms that receive an initial allowance to other firms that applied for a patent but received an initial rejection. Empirically, most patent applications receive an initial decision within three years of being filed (see Appendix Figure E.1). While some applications that are initially rejected receive a patent grant relatively quickly, the modal application that is initially rejected is never granted a patent (see Appendix Figure E.2).

Because our empirical strategy will contrast firms whose applications are initially allowed to those whose applications are initially rejected, having some sense of what predicts initial allowance decisions is useful. Table 3 reports least squares estimates of the probability of an initial allowance as a function of firm characteristics in the year of application. Column (1a) shows that predicting initial allowances is surprisingly difficult. Applications from firms with more W2 employees are somewhat less likely to be initially allowed, as are those from firms with higher value added per worker. Jointly, the covariates are statistically significant. Column (1b) adds art unit by application year fixed effects that control for technology-specific changes over time. This simple addition renders all baseline covariates statistically insignificant both individually and jointly, which provides some assurance that initial patent decisions are not strongly dependent on baseline firm performance. Given this empirical evidence, we proceed by assuming that any remaining selection is on time-invariant firm characteristics that can be captured by firm fixed effects.

A separate concern has to do with whether initial allowances are best thought of as idiosyncratic or market-level shocks. Seminal work by Jaffe, Trajtenberg, and Henderson (1993) demonstrated that patent citations are highly localized geographically. To test whether initial allowances are also geographically clustered, we fit linear random effects models to the initial allowance decision. Appendix Table E.1 reports intraclass correlations at various levels of geography before and after subtracting off art unit by application year mean allowance rates. In either case, the within-state correlation is estimated to be zero, while the correlation within five-digit ZIP codes is quite low (0.07-0.10) and statistically indistinguishable from zero. These findings indicate that initial allowances are best thought of as truly idiosyncratic firm-specific shocks that are unlikely to elicit market-wide wage responses.

5   Detecting Valuable Patents

The value distribution of granted patents is heavily skewed (Pakes 1986), which suggests that low-value patent applications—if granted—are unlikely to generate meaningful shifts in firm outcomes. Constructing a measure of the ex-ante value of patent applications enables us to focus our analysis on patent applications that are likely to induce changes in firm behavior.

A variety of metrics have been proposed as measures of the value of granted patents, including forward patent citations (Trajtenberg 1990), 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 market returns (Kogan et al. 2017). These value measures encounter three challenges in our empirical context. First, these measures are only defined for granted patents, whereas we would like to take advantage of data on patent applications, including those that are ultimately unsuccessful. Second, most of these measures arguably correspond to a measure of social value—or social spillovers, in the sense of social value minus private value—whereas we

are more interested in measuring firms' private value of a patent. This issue arises most sharply with forward patent citations, which are typically used as a measure of spillovers (e.g., Bloom, Schankerman, and Van Reenen 2013). Third, all of these measures are defined ex-post: citations, renewals, reassignments, and litigation are often measured many years after the initial patent award. But in our context—as in Kogan et al. (2017)—what is arguably more relevant is the expected private value of the patent at the time of the patent application or patent grant.

To this end, we build on the recent analysis of Kogan et al. (2017) (henceforth, KPSS), who measure the high-frequency response of stock prices around the date of patent grant announcements to estimate the value of patent grants that are awarded to publicly traded companies. We estimate a simple statistical model designed to extrapolate their estimates to non-publicly traded companies and to non-granted patent applications in our analysis sample.

We model the KPSS patent value ξ_j for each firm-patent application j in our data as obeying the following conditional mean restriction:

$$E\left[\xi_j | X_j, A_j\right] = \exp\left(X_j'\delta + v_{A_j}\right),$$

where X_j denotes a vector of baseline firm and patent application covariates and A_j denotes the art unit to which the application was assigned. The exponential functional form underlying this specification is designed to accommodate the fact that the KPSS values are non-negative and heavily skewed. Because we have, on average, only 2.3 applications with non-missing ξ_j per art unit, some penalization is required to avoid overfitting. Accordingly, we treat the art unit effects $\{v_a\}$ as *i.i.d.* draws from a normal distribution with unknown variance σ_v^2 rather than fixed parameters to be estimated. The model is fit via a random effects Poisson maximum likelihood procedure. As described in Appendix C, this procedure exploits the conditional mean restriction $E\left[\xi_a | X_a\right] = \int \exp\left(X_a'\delta + v\right)\omega_a\left(v\right)dv$ where ξ_a is the vector of KPSS values in an art unit a, X_a is the corresponding vector of baseline application and firm predictors, and $\omega_a\left(v\right)$ is the posterior distribution of v_a given the observed data (ξ_a, X_a).

Table 4 reports the Poisson parameter estimates. Applications submitted to more countries ("patent family size") tend to be of higher value, as do applications with more claims and applications submitted by firms with larger revenues.[9] We also document substantial variability of patent value across art units: a standard deviation increase in the art unit random effect is estimated to raise mean patent values by 127 log points. This variability finding is of interest in its own right as it suggests that patent decisions involve much higher stakes in some USPTO art units than others.

---

[9]The number of countries to which an application was submitted, often referred to as patent "family size," is defined as a set of patent applications filed with different patenting authorities (e.g., the US, Europe, Japan) that refer to the same invention; work starting with Putnam (1996) has argued that firms should be willing to file more privately valuable patents in a larger number of countries. Patents list "claims" over specific pieces of intellectual property, and work starting with Lanjouw and Schankerman (2001), has argued that patents with a larger number of claims may be more privately valuable. See Appendix A for details on both these measures.

We use our estimates of the parameters (δ, σ_v) to compute Empirical Bayes predictions $\hat{\xi}_j$ of ξ_j for every patent application in our analysis sample, including those that lack a KPSS value either because the application is assigned to a privately held firm, or because the application is never granted a patent.[10] Empirically, these predictions are highly accurate: a least squares fit of ξ_j to $\hat{\xi}_j$ yields a slope of 1.12 and an R^2 of 68%. Figure 1 shows that binned average KPSS values track the Empirical Bayes predictions very closely. Appendix Table E.2 lists mean predicted values by subject matter area.

The ultimate test of $\hat{\xi}_j$ is whether it predicts treatment effect heterogeneity: that is, do allowances of patent applications of higher predicted value result in larger shifts in firm outcomes? To investigate this question, we fit a series of interacted difference-in-differences models of the following form:

$$Y_{jt} = \alpha_j + \kappa_{t,k(j)} + Post_{jt} \cdot \left[\sum_{b=1}^{5} s_b\left(\hat{\xi}_j\right) \cdot \left(\tilde{\psi}_b + \tilde{\tau}_b \cdot IA_j\right) \right] + r_{jt} \tag{5}$$

where Y_{jt} is an outcome for firm (EIN) j in year t, α_j are firm fixed effects, and $\kappa_{t,k(j)}$ are calendar year fixed effects that vary by art-unit/application year cell $k(j)$. The variable $Post_{jt}$ is an indicator for having received an initial patent decision, IA_j is an indicator for whether the patent application is initially allowed, and $\{s_b(.)\}_{b=1}^{5}$ is a set of basis functions defining a natural cubic spline with five knots.[11] Intuitively, this specification compares initially allowed and initially rejected applications in the same art unit by application year cell, before and after the date of the initial decision. The spline interactions allow the effects of an initial allowance to vary flexibly with the predicted patent value $\hat{\xi}_j$.

Of primary interest is the "dose-response" function $d(x; \tilde{\tau}) \equiv \sum_{b=1}^{5} s_b(x) \tilde{\tau}_b$, which gives the effect of an initial allowance for a patent with predicted value x. Figure 2 plots our estimates of this function for a grid of values x when Y_{jt} is either operating surplus per worker or wage bill per worker. In both cases, we find evidence of an S-shaped response: impacts of initial allowances on both wages and operating surplus are small and statistically insignificant at low predicted value levels, corroborating both the exclusion and random assignment assumptions underlying our research design. Patents with ex-ante predicted patent values above $5 million in 1982 USD—roughly the 80th percentile of the predicted value distribution—have larger, statistically significant treatment effects that increase rapidly before stabilizing at values near $12 million in 1982 USD.[12]

Given the S-shaped pattern of treatment effect heterogeneity documented in Figure 2, our empirical analysis pools the bottom four quintiles together and focuses on estimating the impacts of patents in the top quintile of

---

[10]In cases where no valid KPSS values are present in the entire art unit, we form our prediction by imputing an art unit random effect of zero.

[11]The natural cubic spline is a cubic b-spline that imposes continuous second derivatives everywhere but allows the third derivative to jump at the knots (see Hastie, Tibshirani, and Friedman 2016 for discussion). Following Harrell (2001), we space knots equally at the 5th, 27.5[th], 50th, 72.5[th], and 95th percentiles of the distribution of patent values. The spline is constrained to be linear below the 5th and above the 95th percentiles.

[12]We reference 1982 dollars because those are the units used by KPSS.

ex-ante predicted patent value. Reassuringly, columns (2a) and (2b) of Table 3 show that initial allowances are equally difficult to predict with baseline characteristics within the top quintile of predicted value, especially after art unit by application year fixed effects have been included. Likewise, columns (3a)-(4b) of Appendix Table E.1 show that among top-quintile applications, initial allowances continue not to exhibit spatial correlation.

6 Reduced Form Estimates

The treatment effect heterogeneity documented in Figure 2 demonstrates that firms experience economically and statistically significant increases in profitability and wages when valuable patent applications are allowed. However, a natural concern is that these findings could reflect pre-existing trends rather than causal effects of the patent decisions themselves. In order to investigate this concern, we estimate a series of "event study" specifications of the following form:

$$
\begin{aligned}
Y_{jt} \;=\; & \alpha_j + \kappa_{t,k(j)} + Q5_j \cdot \left[\sum_{m \in \mathcal{M}} D_{jt}^m \cdot (\psi_{5,m} + \tau_{5,m} \cdot IA_j) \right] \\
& + (1 - Q5_j) \cdot \left[\sum_{m \in \mathcal{M}} D_{jt}^m \cdot (\psi_{<5,m} + \tau_{<5,m} \cdot IA_j) \right] + r_{jt}
\end{aligned}
\tag{6}
$$

where $Q5_j$ is an indicator for the firm's patent application being in the top quintile of predicted ex-ante value, D_{jt}^m is an indicator for firm j's decision having occurred m years ago, and the set $\mathcal{M} = \{-5, -4, -3, -2, 0, 1, 2, 3, 4, 5\}$ defines the five-year horizon over which we study dynamics.[13] The coefficients $\{\hat{\tau}_{5,m}, \hat{\tau}_{<5,m}\}_{m \in \mathcal{M}}$ summarize the differential trajectory of mean outcomes for initially allowed and initially rejected firms by time relative to the initial decision for top-quintile and lower-quintile value observations, respectively.

Figure 3 plots coefficients from equation (6) for our main firm outcome variable, operating surplus. The estimated coefficients illustrate that, among firms with patent applications in the top quintile of the predicted value distribution, firms whose applications are initially allowed exhibit similar trends in gross surplus per worker to those whose applications are initially rejected in the years prior to the initial decision. However, surplus per worker rises differentially for allowed firms in the wake of an initial allowance, and remain elevated afterwards. Firms with lower predicted value applications, by contrast, exhibit no detectable response of surplus per worker to an initial allowance. Figure 4 documents similar patterns in our main worker outcome variable, wage bill per worker. As expected, the wage response to an initial allowance is muted relative to the surplus response; the ratio of these two impacts provides a crude estimate of the pass-through coefficient π of roughly one-third.

While wages and gross surplus respond rather immediately to top-quintile initial allowances, Figure 5 reveals

---

[13]We "bin" the endpoint dummies so that D_{jt}^5 is an indicator for the decision having occurred five or more years ago and D_{jt}^{-5} is an indicator for the decision being five or more years in the future.

that firm size (as measured by the log number of employees) responds more slowly in response to a patent allowance, taking roughly three years to scale to its new level. The fact that earnings impacts remain stable over this horizon casts doubt on the possibility that the impacts in Figure 3 are driven primarily by an increase in hours worked (which we cannot observe in tax data) rather than an increase in hourly wages. The nearly immediate response of operating surplus and wages to initial allowances may signal that our panel of relatively small innovative firms was initially credit constrained. Evidence from Farre-Mensa, Hegde, and Ljungqvist (2017), who document that patent grants are strongly predictive of access to venture capital financing, corroborates this view. Of relevance to our analysis is that access to venture capital and other forms of financing is a plausible additional channel through which patent decisions could quickly affect the marginal revenue product of labor and consequently worker wages.

Finally, Figure 6 documents that an initial allowance raises the probability of having the patent application granted by roughly 50% in the year after the decision, with gradual declines afterwards. The probability of receiving a patent grant jumps by less than 100% for two reasons. First, some initially allowed applications are not pursued by applicants, possibly because the assignee went out of business while awaiting the initial decision, or because the applicant learned new information since filing which led them to believe that the patent was not commercially valuable. Second, as described in Section 4, many initially denied applications reapply and eventually have their applications allowed. Our estimates in Figure 6 suggest that the initial impact of initial allowances on patent grants is somewhat smaller for higher-value patents, more of which would be approved shortly after a rejection; a pooled difference-in-difference estimate of the impact on the grant probability of high-value patents is approximately one third. Hence, the impact of high-value patent grants on firm outcomes is likely to be roughly three times the impact of an initial allowance on firm outcomes, though it is possible that allowances influence firm outcomes independent of grant status if allowances relieve credit constraints before a patent has actually been granted. In what follows, we continue to report the reduced form impacts of allowances as our ultimate goal is to instrument for operating surplus rather than for patent grants.

6.1 Impacts on Firm Averages

Table 5 pools pre- and post-application years and quantifies the average effects displayed in the event study figures by fitting simplified difference-in-differences models of the following form:

$$Y_{jt} = \alpha_j + \kappa_{t,k(j)} + Q5_j \cdot Post_{jt} \cdot (\psi_5 + \tau_5 \cdot IA_j) \tag{7}$$
$$+ (1 - Q5_j) \cdot Post_{jt} \cdot (\psi_{<5} + \tau_{<5} \cdot IA_j) + r_{jt}.$$

The parameters reported in Table 5 are τ_5 and $\tau_{<5}$, which respectively govern the effects of top-quintile and lower-quintile value patents being initially allowed.

Column 1 of Table 5 documents that initial allowances have no effect on the probability of firm survival, as proxied by the presence of at least one W2 employee. Given this result, the remainder of the columns in this table focus on outcomes conditional on firm survival as measured by the presence of at least one W2 employee (hence the smaller sample sizes in subsequent columns). Column 2 of Table 5 reports the impact of an initial allowance on firm size, as measured by the number of W2 employees at the firm. Having a top-quintile patent allowed leads the firm to expand by 19 workers on average, which is roughly 30% of the mean firm size of top-quintile value firms in the application year. Notably, initial allowances of patents with lower predicted value have no detectable impact on firm survival, firm size, or any other outcome we examine; these results suggest that differential trends for initially allowed and initially rejected patents are unlikely to confound our analysis.

An allowance of a high-value patent application is associated with roughly $37,000 in additional revenue per worker (Column 3 of Table 5) and roughly $16,000 in value added per worker (Column 4 of Table 5). EBITD per worker rises by roughly $9,000 (Column 5 of Table 5), which we interpret as income to firm owners, while wage bill per employee rises by roughly $4,000 (Column 6 of Table 5). Our gross surplus measure, which sums EBITD and wage bill, rises by $12,000 per worker (Column 7 of Table 5).[14] As described in Section 3, we interpret our estimated effects on gross surplus as the impact on total operating cash flow at the firm. Our central interest in this paper is on how this gross surplus measure is divided between workers and firm owners.

For reference, Panel B of Table 5 documents impacts on $\sinh^{-1}(.)$ transformations of the dependent variables. This transformation is a concave function that reduces skew but accepts zero and negative values. Impacts on transformed variables can be interpreted as being analogous to percentage impacts.[15] Another interpretation, derived in Appendix D, is that impacts on transformed variables place less weight on impacts in the upper quantiles of the conditional distribution of outcomes. While the estimated impacts on $\sinh^{-1}(.)$ transformed firm size and gross surplus remain statistically distinguishable from zero at conventional levels, impacts on revenue, EBITD per worker, and value added per worker fall to statistical insignificance. This weakening of impacts suggests that impacts on revenue, EBITD, and value added per worker are largest in the upper quantiles of the distribution.

Table 5 also reports impacts on various measures of labor compensation. A successful top-quintile patent application is associated with an increase in firm-level deductions for labor-related expenses of around $4,000 (Column 8 of Table 5), roughly comparable to what we found for wage and salary compensation based on W2 wage bills. On the other hand, pooling W2 earnings with 1099 earnings yields an impact of only $2,800 per worker (Column 9 of Table 5). The $\sinh^{-1}(.)$ transformations of these compensation measures yield impacts of roughly 3% and 7%, respectively, which can be compared to the corresponding impact on W2 wages per worker

---

[14]The sum of the per-worker impacts on EBITD and wage bill does not exactly match the impact on surplus per worker because the variables are winsorized separately.

[15]A Taylor approximation gives $d\sinh^{-1}(z) = (1+z^2)^{-1/2}\,dz$ which, for large z, is approximately equal to the percent change $\frac{dz}{z}$. In practice, we find impacts on $\sinh^{-1}(.)$ and logarithmic transformations of the strictly positive outcomes to be nearly identical.

of roughly 8%. Together, these results suggest that 1099 compensation is, if anything, less responsive to shocks than W2 wages and salaries.

Finally, the last column of Table 5 reports impacts on a measure of the average individual income tax burden per worker.[16] An initial allowance of a high-value patent is estimated to yield $840 of additional tax revenue per worker. Although this figure is statistically indistinguishable from zero, the point estimate implies an effective marginal tax rate of 23% on the $3,600 of extra W2 earnings reported in Column 6 of Table 5, which is roughly the average US marginal tax rate found in TAXSIM (Feenberg and Coutts, 1993) over our sample period.[17] A $\sinh^{-1}(.)$ transformation of tax burden per worker indicates that an initial allowance of a high-value patent raises tax revenue per worker by 8%—roughly the same proportional impact as found on W2 earnings per worker. This finding suggests the presence of an important fiscal externality between corporate tax treatment of innovation and income tax revenue.

6.2 Impacts on Workforce Composition

One difficulty with interpreting impacts on firm-level aggregates is that firms may alter the skill mix of their employees in response to shocks, in which case changes in wages could simply reflect compositional changes rather than changes in the compensation of similar employees. We investigate the possibility of such compositional changes in Table 6.

Columns 1 and 2 of Table 6 reveal that neither the share of employees who are women nor the share of employees who are inventors changes appreciably in response to an allowance. We also find little evidence that the quality of new hires ("entrants"), as proxied by their earnings in the year prior to hiring (Column 3 of Table 6), rises in response to an initial allowance. Likewise, the earnings of those workers who choose to separate from the firm appear to be unaffected by the allowance (Column 4 of Table 6).

Examining "firm stayers" who were present in the year of application and choose to remain with the firm provides a different window into potential changes in workforce composition. An initial allowance increases the number of workers present in the year of application who remain employed at the firm by roughly 12% (Column 5 of Table 6). Notably, the share of stayers in total employment does not fall appreciably (Column 6 of Table 6), indicating that the firm size responses documented in Table 5 are attributable, in part, to improvements in worker retention. However, we find no appreciable effect on the application year earnings of stayers (Column 7 of Table 6), suggesting that this increased retention does not induce a significant amount of selection. Finally the average age of W2 employees drops by roughly a year in response to a valuable patent allowance (Column 8 of Table 6), which

---

[16]Our measure, which is the main tax variable in the databank, captures "tentative" tax burden before accounting for the Alternative Minimum Tax.

[17]See `http://users.nber.org/~taxsim/allyup/ally.html` for annual estimates.

is in keeping with our finding that firms grow in response to valuable allowances and the fact that job mobility declines with age (Farber 1994). Taken together, these results indicate that skill upgrading responses, if present, are likely to be small and that modest skill downgrading (through age declines) is an equally likely possibility.

6.3 Impacts on Within-Firm Inequality

Table 7 analyzes the impact of initial allowances on various measures of within-firm inequality. Consistent with the literature on gender differences in rent-sharing (e.g., Black and Strahan 2001; Card, Cardoso, and Kline 2016), we find that initial allowances exacerbate the gender earnings gap. While male earnings rise by roughly $6,000 (or roughly 9%; Column 1 of Table 7) in response to a valuable patent allowance, female earnings appear unresponsive to initial allowances (Column 2 of Table 7). Focusing on firms that employ both genders, we find that the gender earnings gap increases by roughly $7,000 in response to a valuable initial allowance, or roughly 10% (Column 3 of Table 7). While we cannot rule out that patent allowances generate modest earnings increases for women, the gender disparities we estimate are larger than has been found in previous work. There are several potential explanations for why we might estimate a larger discrepancy in our setting. First, negotiations over salary and promotion may be more important at innovative firms, which could place women at a disadvantage (Babcock and Laschever 2003; Bowles, Babcock, and Lai 2007). Second, women at innovative firms may be differentially employed in occupations where pay is tied less tightly to performance (relative to occupations of male employees at the same firms). Third, statistical discrimination against women may be more prevalent in innovative sectors, making innovative firms particularly reluctant to promote talented female employees for fear of revealing their quality and stoking costly labor market competition (Milgrom and Oster 1987).

The earnings gap between inventors and non-inventors also widens in response to an initial allowance. Column 4 of Table 7 shows that the earnings of inventors rise by roughly $17,000 in response to an initial allowance; however, this effect is only distinguishable from zero at the 10% level. The earnings of non-inventors rise by only around $2,000 in response to an allowance, though the impacts on the $\sinh^{-1}(.)$ transformation suggest an approximately 6% rise in wages that is distinguishable from zero at the 5% level (Column 5 of Table 7). Focusing on firms that employ both inventors and non-inventors, we find that the inventor-non inventor earnings gap increases by roughly $14,000 in response to a valuable initial allowance, or roughly 12% (Column 6 of Table 7). The gender and inventor gaps are overlapping, but not identical phenomena. Column 7 of Table 7 shows that the earnings of non-inventor males rise by roughly $3,500 – less than all males, but more than all non-inventors.

Finally, to provide a composite measure of within-firm earnings inequality, we break workers in each firm-year with at least four W2s into quartiles based on their annual earnings. We find no effect of an initial allowance on the average earnings of bottom quartile workers (Column 8 of Table 7), but find that mean earnings of top-

quartile workers rises by roughly $8,000 per worker (Column 9 of Table 7). The pay gap between top and bottom quartile workers rises by roughly the same amount (Column 10 of Table 7). This finding suggests that standard log-additive econometric models of firm wage effects of the sort pioneered by Abowd, Kramarz, and Margolis (1999) may miss an important role for firms in generating within-firm wage dispersion in addition to the usual between-firm component that has been emphasized in the past literature (e.g., Card, Heining, and Kline 2013; Bloom, Schankerman, and Van Reenen 2013; Card et al. forthcoming).

6.4   Impacts on Earnings by Timing of Worker Entry and Exit

Our results in Section 6.2 suggested that initial allowances are not associated with major changes in workforce composition. However, an alternative way to hold constant the quality of the workforce is to study the impact of a patent allowance on the earnings of a fixed cohort of workers. As we illustrate below, parsing impacts on workers separately by their time of entry and exit to the firm also provides a way to link our empirical estimates more closely to the predictions of various theoretical models of wage setting.

Column 1 of Table 8 documents that the average earnings of the cohort of workers present in the year of the patent application rise by roughly $4,000 or about 6% in response to an initial allowance. These effects are concentrated in the subset of the cohort that remains with the applicant firm ("stayers"), whose earnings are estimated to rise by $8,000 (or 10%) per year in response to an initial allowance (Column 2 of Table 8). Members of the application cohort who leave the firm, by contrast, have earnings that fall insignificantly in response to an initial allowance (Column 3 of Table 8). The concentration of earnings effects on stayers casts some doubt on reputational (or "career concerns") explanations for firm-specific wage fluctuations (Harris and Holmström 1982; Gibbons and Murphy 1992; Holmström 1999), as firm leavers appear to be unable to transport their patent-induced wage gains to new employers.

The monopsony model of Section 2 was predicated on the existence of a common wage for new hires and incumbent workers. Interpreted through the lens of that model, it is rather surprising that we find an economically small and statistically insignificant effect of initial allowances on the average earnings of entrants (Column 4 of Table 8). Given our findings in Section 6.2 that the composition of entrants does not seem to have changed in response to initial allowances, the discrepancy between our measured impacts of initial allowances on the earnings of entrants and on the earnings of firm stayers suggests that the order in which workers are hired plays an independent role in the transmission of firm shocks to wages. This order dependence may reflect performance pay, a possibility to which we return later.

As mentioned in Section 2, some dynamic models (e.g., Postel-Vinay and Robin 2002) can generate a drop in entry wages in response to a firm productivity increase because wage growth rates increase. Such an elevation of

growth rates should eventually impact earnings levels. However, Columns 5 and 6 of Table 8 reveal wrong-signed and generally insignificant impacts of initial allowances on the earnings of workers hired within the last three years, or even among all workers who joined the firm after the year of application. A shift in growth rates, in conjunction with stable entry wages, should also lead to an escalating pattern of pooled wage impacts. However, we saw in Figure 4 that wage impacts are roughly stable after the initial decision. Hence, we conclude there is no evidence of a permanent impact on earnings growth rates.[18]

Columns 7 through 9 of Table 8 adjust for possible compositional changes by subtracting from the various earnings measures an average earnings level of the same group of workers in previous years, which adjusts for any time-invariant heterogeneity in worker quality. Column 7 shows that subtracting the average application year earnings of the firm stayers has little effect on the estimates. The estimates in Columns 8 and 9 remain statistically insignificant and continue to hover near zero, suggesting that these other groups' earnings are relatively insensitive to the patent decision.

Finally, Figure 7 reports impacts of high-value initial allowances on the within-firm earnings distribution for all workers, firm stayers, and entrants, respectively. Panels A and C reveal that the earnings impacts of patent allowances are strikingly concentrated among workers in the top quartile of the contemporaneous earnings distribution. While this is consistent with recent concerns (e.g., Bebchuk and Fried 2003, 2006; Kuhnen and Zwiebel 2008) that executive pay is set in an inefficient manner that rewards—among other factors, luck (Bertrand and Mullainathan 2001; Piketty, Saez, and Stantcheva 2014)—we cannot rule out that these responses in fact reflect efficient contracts (a point emphasized by Edmans and Gabaix 2009). Consistent with our pooled results, Panels B and D reveal that impacts on the average earnings of new hires ("entrants") are uniformly statistically insignificant and concentrated near zero, regardless of the contemporaneous earnings quartile being joined. By contrast, the earnings gains of firm stayers appear to be broadly shared, with the proportional impacts in Panel C being roughly constant across quartiles defined by the worker's position in the firm's earnings distribution *at the time of the patent application*. This pattern strongly suggests that our earlier findings do not solely represent—for example—the capture of rents by CEOs or other top executives.

7   Pass-Through Estimates

Table 9 reports "rent-sharing" specifications that relate earnings outcomes to surplus per worker. Regressing average wage bill per worker on surplus per worker, together with our standard set of (firm and art-unit by application year by calendar year) fixed effects yields an estimated pass-through coefficient π of 0.16. Instrumenting

---

[18]We have also directly computed impacts on earnings growth rates for workers hired within the last three years, but this led to highly imprecise estimates. Specifically, we estimate an impact of negative one percentage point on the three-year growth rate of the earnings of new hires, with a standard error of seven percentage points.

gross surplus with the interaction of a post-decision indicator and an indicator for the application being initially allowed increases the estimated coefficient to 0.29, implying that workers capture 29 cents of each additional dollar of operating surplus. For comparison with the prior literature, we also report specifications taking a $\sinh^{-1}(.)$ transformation of both surplus and wages, which captures an approximate elasticity. While OLS estimation yields a pass-through elasticity of only 0.04, IV yields an elasticity of 0.19 that is statistically distinguishable from the OLS estimate.[19]

Column 2 of Table 9 restricts attention to firm stayers, who were present in the year of application. OLS estimates indicate that stayer earnings are more sensitive to surplus fluctuations in levels (relative to the sample of all workers), but the elasticity is the same as was found for the average earnings of all workers (0.04). Instrumenting operating surplus changes this conclusion dramatically: stayers are estimated to capture 60 cents of every dollar of surplus, with a corresponding elasticity of 0.23. Column 3 adjusts stayer earnings for potential changes in workforce composition by subtracting off their earnings in the application year, which should difference out any selection on time invariant worker skills. As expected given our results in Section 6.2, this adjustment has minor effects on the results—lowering, for instance, the instrumented pass-through of surplus to earnings from 60 cents to 50 cents on the dollar. Finally, Column 4 shows that these results are not driven exclusively by workers listed as inventors on patent applications: the instrumented value of π among non-inventor stayers is 0.49.

In sum, we find that workers, particularly those who were present in the year of application, capture a large fraction of operating surplus. In elasticity terms, our estimates are larger than the bulk of recent studies reviewed by Card, Cardoso, and Kline (2016) and fall closer to the earlier estimates of Abowd and Lemieux (1993) and Van Reenen (1996). This difference could reflect differences between our sample of innovative firms and firms in less innovative sectors. However, we suspect that some of this divergence is also attributable to the fact that the studies of Abowd and Lemieux (1993) and Van Reenen (1996) utilized external instruments, while the recent literature reviewed by reviewed by Card, Cardoso, and Kline (2016) tends to rely on statistical assumptions regarding value added per worker to identify shocks. Notably, our OLS estimates are in line with this recent literature, which typically finds pass-through elasticities somewhat below 0.1.

8 Labor Supply Estimates and Model-Based Interpretations

The wage posting model presented in Section 2 interprets earnings responses to firm-specific shocks in terms of movements along a firm-specific labor supply curve. It is therefore of interest to ask what sorts of parameters are necessary to rationalize our findings through the lens of that model.

---

[19]In theory, this larger IV estimate could reflect a tendency for wages to respond to lower frequency fluctuations in surplus (Guiso, Pistaferri, and Schivardi 2005). We investigate this possibility in Appendix Table E.3, which shows that using three-year averages of surplus yields very small increases in estimates of π.

Table 10 uses the estimated impacts of initial patent decisions on wages and employment to estimate labor supply elasticities to the firm. A naive least squares fit of $\sinh^{-1}$ (employees) to $\sinh^{-1}$ (wage bill per worker) with our standard battery of (firm and art-unit by application year by calendar year) fixed effects yields a labor supply elasticity of only 0.19, which would indicate that firms in our analysis sample have very substantial market power. Instrumenting with the allowance decision raises the estimated elasticity to 2.7, which is statistically distinguishable from the OLS estimate. This estimate is towards the upper end of the labor supply estimates reported in Manning (2011), but a bit lower than the benchmark considered in Card et al. (forthcoming). Recall that in the monopsony model workers are paid a fraction $\frac{\eta_j(q)}{1+\eta_j(q)}$ of their marginal revenue product. Ignoring elasticity heterogeneity for the moment, our labor supply estimate implies that wages are marked down roughly 27% from competitive levels.

Another useful benchmark comes from noting that equation (1) implies that the ratio of offered wages to reservation wages for a worker of quality q can be written:

$$\frac{w_j(q)}{b_j(q)} = \frac{\eta_j(q)}{\eta_j(q) - \beta}.$$

Using our pooled rent-sharing estimate of $\pi = .29$, we have $\beta = \frac{.29/\varepsilon}{1 - .29/\varepsilon}$. Recent work has used values of ε ranging from 4.5 (Suárez Serrato and Zidar 2016) to 7 (Coibion, Gorodnichenko, and Wieland 2012), which suggests that 0.5–0.6 is a reasonable range for β. With $\eta_j(q) = 2.7$, offered wages hover roughly 22%–29% above the reservation wage. This range is quite close to Manning's (2011) estimate in the British Household Panel Survey that the mean expected wage is 21 log points above the reservation wage—a finding that is challenging for standard search theoretic models to match (Hornstein, Krusell, and Violante, 2011).

While the monopsonistic wage posting model yields a plausible interpretation of the pooled impacts on earnings and firm size, it has a more difficult time rationalizing heterogeneity across groups. All else equal, the model predicts that subgroups with greater earnings pass-through should have larger labor supply elasticities. However, Column 2 of Table 10 finds that men exhibit a *below-average* supply elasticity, which contradicts our earlier finding in Section 6.3 that male wages are more responsive to initial allowances than are female wages. Notably, this finding also directly contradicts a longstanding presumption in the monopsony literature dating back to Robinson (1933) that men have higher supply elasticities than women. Column 3 of Table 10 generates another puzzle: non-inventors have an above-average supply elasticity, which is at odds with our finding that inventor wages are more responsive to initial allowances than are the wages of non-inventors.

Perhaps most troubling for the monopsony model is the heterogeneity in pass-through between incumbent workers and new hires. Because the wages of new hires are unresponsive to patent allowances, we cannot reject that the entry market is perfectly competitive. By contrast, Column 4 of Table 10 reveals that, among firm stayers,

26

the wage elasticity of labor supply falls to approximately 1, or to 1.5 if we adjust the wages of firm stayers for selection by differencing off their application year earnings (Column 5 of Table 10). These below-average supply elasticities contradict the greater wage responsiveness of firm stayers. Moreover, because the number of stayers adjusts only via separations, these lower elasticities suggest that separations are less wage elastic than new hires. As noted by Manning (2011), large differences between separation and recruitment elasticities are difficult to rationalize in a wage-posting framework as wage-induced separations from one firm are wage-induced accessions to another.

A plausible reconciliation of these facts is that firms may tie pay and promotion decisions to patent allowances when those allowances convey information about difficult-to-monitor activities at the firm (Mirrlees 1976; Holmström 1979, 1989). This possibility provides a compelling explanation for the failure of entry wages to respond to patent decisions because allowances cannot signal effort on the part of new hires who were not involved in the patent application process. Likewise, it is natural to expect that allowances are more informative about the performance of inventors than the performance of non-inventors, which rationalizes the greater observed pass-through to inventor wages. The greater pass-through to male wages relative to female wages could reflect at least two factors: given the substantial occupational segregation found in US labor markets (Goldin 2014), it could be that women are concentrated in occupations involving easier-to-monitor tasks that do not require performance pay; alternatively, the failure of female wages to respond to patent allowances could reflect gender discrimination in the ways that performance is rewarded (e.g., Milgrom and Oster 1987). Distinguishing between these interpretations is an interesting avenue for future research. To the extent that performance pay considerations are at play, the wage responses uncovered here likely reflect "risk-sharing" in addition to rent-sharing; that is, while the earnings gains of the employees at initially allowed firms almost certainly constitute a rent ex-post, they do not necessarily signal the presence of ex-ante rents at the time of the hiring decision. Under either the rent- or risk-sharing interpretations, however, our estimates provide a valid measure of the contribution of firm shocks to wage fluctuations.

9   Conclusion

This paper analyzes how patent-induced shocks to labor productivity propagate into worker earnings using a new linkage of US patent applications to US business and worker tax records. Our baseline estimates suggest that every patent-induced dollar of gross operating surplus yields 29 cents of additional worker earnings, with an approximate pass-through elasticity of 0.2. These estimates provide some of the first evidence that truly idiosyncratic variability in firm performance is an important causal determinant of worker pay. Given that firm productivity is highly variable and persistent (Luttmer 2007; Foster, Haltiwanger, and Syverson 2008), it is plausible that firm-specific shocks contribute substantially to permanent earnings inequality among identically skilled workers. For

example, Foster, Haltiwanger, and Syverson (2008) find in the Census of Manufacturing that the standard deviation of annual shocks to total factor productivity is on the order of 10%. Taken at face value, our estimated pass-through elasticity of 0.2 implies that a one standard deviation firm productivity shock would persistently raise wages by around 2%.

We document several significant sources of heterogeneity in the pass-through of patent-induced shocks to workers that give rise to increases in within-firm inequality. First, we document that the earnings of male employees strongly respond to patent decisions, whereas the earnings of female employees show little response. Second, while the earnings of both inventors and non-inventors respond to patent decisions, earnings of inventors are more responsive. Finally, in the full sample of employees, earnings impacts are strongly concentrated among employees in the top quartile of the within-firm earnings distribution. However, among firm stayers, earnings seem to rise throughout the distribution, with proportional impacts that are roughly constant across quartiles. These findings strongly suggest that firms play an important role in generating earnings inequality not only *across* but also *within* workplaces.

Intriguingly, we also uncover evidence that the wages of incumbent workers respond more strongly to patent decisions than do wages of new workers: incumbent workers—including non-inventors—gain 50 cents or more of each dollar of additional operating surplus, whereas the earnings of new hires appear insensitive to patent decisions. This differential response cautions strongly against interpreting the firm-specific component of earnings variability as an omnibus indicator of market power. Rather, our results suggest that the wage fluctuations that we document in part reflect risk-sharing behavior, which is a potentially optimal response to moral hazard considerations (Holmström 1979). Since many employment relationships are likely to involve a mix of both risk- and rent-sharing behavior (e.g., as in Lamadon 2016), further research is needed to isolate robust signs of inefficiency applicable to a wide range of models.

From a tax policy perspective, an implication of our findings is that patents influence the revenue raised from both business and individual income taxes. Consequently, so-called "patent box" proposals, which are designed to exempt the rents associated with patent grants from business taxes, are likely also to impact the revenue collected from individual income taxes.[20] Specifically, our findings suggest that the direct fiscal costs of patent boxes in terms of decreased corporate tax revenue are likely to be partially offset by increases in the revenue raised from individual income taxes. The potential for this type of fiscal externality has so far been ignored in discussions of patent box policies (e.g, Griffith, Miller, and O'Connell 2014; Furman 2016) as well as many other business tax

---

[20]Although the specifics of patent box policies vary by country, patent boxes tax income derived from intellectual property at a rate below the typical corporate income tax rate. Patent boxes have been controversial partly due to concerns that they have spurred tax competition effects, with research-intensive companies relocating to take advantage of the policies; for example, US-based Pfizer's attempt to take over UK-based Astra Zeneca was widely perceived as an attempt to take advantage of the UK's patent box policy (Economist 2014; Evans 2014; Houlder 2014).

policies.

References

Abowd, John and Thomas Lemieux, "The Effects of Product Market Competition on Collective Bargaining Agreements: The Case of Foreign Competition in Canada," *Quarterly Journal of Economics*, 1993, *108* (4), 983–1014.

__ **, Francis Kramarz, and David Margolis**, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

__ **, Kevin McKinney, and Nellie Zhao**, "Earnings Inequality and Mobility Trends in the United States: Nationally Representative Estimates from Longitudinally Linked Employer-Employee Data," *Journal of Labor Economics*, forthcoming.

Acemoglu, Daron and William Hawkins, "Search with Multi-Worker Firms," *Theoretical Economics*, 2014, *9* (3), 583–628.

Aghion, Philippe and Jean Tirole, "The Management of Innovation," *Quarterly Journal of Economics*, 1994, *109* (4), 1185–1209.

__ **, Ufuk Akcigit, Ari Hyytinen, and Otto Toivanen**, "Living the American Dream in Finland: The Social Mobility of Inventors," 2017. Working paper.

Babcock, Linda and Sara Laschever, *Women Don't Ask: Negotiation and the Gender Divide*, Princeton, NJ: Princeton University Press, 2003.

Baker, George, Michael Gibbs, and Bengt Holmström, "The Internal Economics of the Firm: Evidence from Personnel Data," *Quarterly Journal of Economics*, 1994, *109* (4), 881–919.

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, "It's Where You Work: Increases in Earnings Dispersion across Establishments and Individuals in the U.S.," *Journal of Labor Economics*, forthcoming.

Bebchuk, Lucian and Jesse Fried, "Executive Compensation as an Agency Problem," *The Journal of Economic Perspectives*, 2003, *17* (3), 71–92.

__ **and** __ , *Pay without Performance: The Unfulfilled Promise of Executive Compensation*, Cambridge, MA: Harvard University Press, 2006.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen, "The Lifecycle of Inventors," 2016. Working paper.

Bender, Stefan, Nicholas Bloom, David Card, John Van Reenen, and Stefanie Wolter, "Management Practices, Workforce Selection and Productivity," *Journal of Labor Economics*, forthcoming.

Bertrand, Marianne and Sendhil Mullainathan, "Are CEOs Rewarded for Luck? The Ones without Principals Are," *Quarterly Journal of Economics*, 2001, *116* (3), 901–932.

Bessen, James, "The Value of U.S. Patents by Owner and Patent Characteristics," *Research Policy*, 2008, *37* (5), 932–945.

Black, Sandra and Philip Strahan, "The Division of the Spoils: Rent-Sharing and Discrimination in a Regulated Industry," *American Economic Review*, 2001, *91* (4), 814–831.

Blanchflower, David, Andrew Oswald, and Peter Sanfey, "Wages, Profits, and Rent-Sharing," *Quarterly Journal of Economics*, 1996, *111* (1), 227–251.

Bloom, Nicholas and John Van Reenen, "Measuring and Explaining Management Practices Across Firms and Countries," *Quarterly Journal of Economics*, 2007, *122* (4), 1351–1408.

_ , **Mark Schankerman, and John Van Reenen**, "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

Bowles, Hannah Riley, Linda Babcock, and Lei Lai, "Social Incentives for Gender Differences in the Propensity to Initiate Negotiations: Sometimes It Does Hurt to Ask," *Organizational Behavior and Human Decision Processes*, 2007, *103* (1), 84–103.

Brown, James and Orley Ashenfelter, "Testing the Efficiency of Employment Contracts," *Journal of Political Economy*, 1986, *94* (3), S40–S87.

Bull, Clive, "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, 1987, *102* (1), 147–159.

Cahuc, Pierre and Etienne Wasmer, "Does Intrafirm Bargaining Matter in the Large Firm's Matching Model?," *Macroeconomic Dynamics*, 2001, *5* (5), 742–747.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women," *Quarterly Journal of Economics*, 2016, *131* (2), 633–686.

_ , _ , **Jörg Heining, and Patrick Kline**, "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, forthcoming.

_ , **Francesco Devicienti, and Agata Maida**, "Rent-Sharing, Holdup, and Wages: Evidence from Matched Panel Data," *Review of Economic Studies*, 2014, *81* (1), 84–111.

_ , **Jörg 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, Julián Messina, and Oskar Nordstrom Skans, "Wage Adjustment and Productivity Shocks," *Economic Journal*, 2016, *126* (595), 1739–1773.

Christofides, Louis and Andrew Oswald, "Real Wage Determination and Rent-Sharing in Collective Bargaining Agreements," *Quarterly Journal of Economics*, 1992, *107* (3), 985–1002.

Cohen, William, Pradeep Ravikumar, and Stephen Fienburg, "A Comparison of String Distance Metrics for Name-Matching Tasks," in "IIWeb" 2003, pp. 73–78.

Coibion, Olivier, Yuriy Gorodnichenko, and Johannes Wieland, "The Optimal Inflation Rate in New Keynesian Models: Should Central Banks Raise Their Inflation Targets in Light of the Zero Lower Bound?," *Review of Economic Studies*, 2012, *79* (4), 1371–1406.

DeBacker, Jason Matthew, Bradley Heim, Shanthi Ramnath, and Justin Ross, "The Impact of State Taxes on Pass-Through Businesses: Evidence from the 2012 Kansas Income Tax Reform," 2016. Working paper.

Economist, "Pfizer and AstraZeneca: A Matter of Trust," *The Economist*, 2014.

Edmans, Alex and Xavier Gabaix, "Is CEO Pay Really Inefficient? A Survey of New Optimal Contracting Theories," *European Financial Management*, 2009, *15* (3), 486–496.

Evans, Peter, "Pfizer Hopes U.K.'s Patent Box Holds Gifts in AstraZeneca Deal," *The Wall Street Journal*, 2014.

Farber, Henry, "The Analysis of Interfirm Worker Mobility," *Journal of Labor Economics*, 1994, *12* (4), 554–593.

Farre-Mensa, Joan, Deepak Hegde, and Alexander Ljungqvist, "What Is a Patent Worth? Evidence from the U.S. Patent 'Lottery'," 2017. NBER working paper no. 23268.

Feenberg, Daniel and Elisabeth Coutts, "An Introduction to the TAXSIM Model," *Journal of Policy Analysis and Management*, 1993, *12* (1), 189–194.

Feigenbaum, James, "Automated Census Record Linking: A Machine Learning Approach," 2016. Working paper.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 2008, *98* (1), 394–425.

Frakes, Michael and Melissa Wasserman, "Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents? Evidence from Micro-Level Application Data," *Review of Economics and Statistics*, 2017, *99* (3), 550–563.

Furman, Jason, "Encouraging Innovation and the Role of Tax Policy," Technical Report, Joint International Tax Policy Forum & Georgetown University Law Center Conference 2016.

Gibbons, Robert and Kevin J. Murphy, "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," *Journal of Political Economy*, 1992, *100* (3), 468–505.

Goldberg, Pinelopi Koujianou and Rebecca Hellerstein, "A Structural Approach to Identifying the Sources of Local Currency Price Stability," *Review of Economic Studies*, 2013, *80* (1), 175–210.

Goldin, Claudia, "A Grand Gender Convergence: Its Last Chapter," *American Economic Review*, 2014, *104* (4), 1091–1119.

Goldschmidt, Deborah and Johannes Schmieder, "The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure," *Quarterly Journal of Economics*, 2017, *132* (3), 1165–1217.

Gorodnichenko, Yuriy and Oleksandr Talavera, "Price Setting in Online Markets: Basic Facts, International Comparisons, and Cross-Border Integration," *American Economic Review*, 2017, *107* (1), 249–282.

Graham, Stuart and Deepak Hegde, "Do Inventors Value Secrecy in Patenting? Evidence from the American Inventor's Protection Act of 1999," 2014. Unpublished USPTO mimeo.

Griffith, Rachel, Helen Miller, and Martin O'Connell, "Ownership of Intellectual Property and Corporate Taxation," *Journal of Public Economics*, 2014, *112*, 12–23.

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

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "The NBER U.S. Patent Citations Data File: Lessons, Insights, and Methodological Tools," 2001. NBER working paper no. 8498.

Hall, Robert and Paul Milgrom, "The Limited Influence of Unemployment on the Wage Bargain," *American Economic Review*, 2008, *98* (4), 1653–1674.

Harhoff, Dietmar, Frederic Scherer, and Katrin Vopel, "Citations, Family Size, Opposition, and the Value of Patent Rights," *Research Policy*, 2003, *32* (8), 1343–1363.

Harrell, Frank, *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*, corrected ed., New York, NY: Springer, 2001.

Harris, Milton and Bengt Holmström, "A Theory of Wage Dynamics," *The Review of Economic Studies*, 1982, *49* (3), 315–333.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., New York, NY: Springer, 2016.

Helpman, Elhanan, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen Redding, "Trade and Inequality: From Theory to Estimation," *Review of Economic Studies*, 2017, *84* (1), 357–405.

Hildreth, Andrew, "Rent-Sharing and Wages: Product Demand or Technology Driven Premia," *Economics of Innovation and New Technology*, 1998, *5* (2-4), 199–226.

Holmström, Bengt, "Moral Hazard and Observability," *The Bell Journal of Economics*, 1979, *10* (1), 74–91.

_ , "Agency Costs and Innovation," *Journal of Economic Behavior & Organization*, 1989, *12* (3), 305–327.

_ , "Managerial Incentive Problems: A Dynamic Perspective," *The Review of Economic Studies*, 1999, *66* (1), 169–182.

Hornstein, Andreas, Per Krusell, and Giovanni Violante, "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *American Economic Review*, 2011, *101* (7), 2873–2898.

Houlder, Vanessa, "Patent Box Holds Key to Pfizer-AZ Deal's Attraction," *The Financial Times*, 2014.

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

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce, "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 1993, *101* (3), 410–442.

Katz, Lawrence and Kevin M. Murphy, "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 1992, *107* (1), 35–78.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological Innovation, Resource Allocation, and Growth," *Quarterly Journal of Economics*, 2017, *132* (2), 665–712.

Kuhnen, Camelia and Jeffrey Zwiebel, "Executive Pay, Hidden Compensation and Managerial Entrenchment," 2008. Rock Center for Corporate Governance working paper no. 16.

Lamadon, Thibaut, "Productivity Shocks, Long-Term Contracts and Earnings Dynamics," 2016. Working paper.

Lanjouw, Jean and Mark Schankerman, "Characteristics of Patent Litigation: A Window on Competition," *RAND Journal of Economics*, 2001, *32* (1), 129–151.

Lemieux, Thomas, W. Bentley MacLeod, and Daniel Parent, "Performance Pay and Wage Inequality," *Quarterly Journal of Economics*, 2009, *124* (1), 1–49.

Lemley, Mark and Bhaven Sampat, "Is the Patent Office a Rubber Stamp?," *Emory Law Journal*, 2008, *58*, 181–203.

_ **and** _ , "Examining Patent Examination," *Stanford Technology Law Review*, 2010, (4), 1–11.

Lentz, Rasmus and Dale Mortensen, "Labor Market Models of Worker and Firm Heterogeneity," *Annual Review of Economics*, 2010, *2*, 577–602.

Lerner, Josh and Julie Wulf, "Innovation and Incentives: Evidence from Corporate R&D," *The Review of Economics and Statistics*, 2007, *89* (4), 634–644.

Levin, Jonathan, "Relational Incentive Contracts," *American Economic Review*, 2003, *93* (3), 835–857.

Luttmer, Erzo, "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, 2007, *122* (3), 1103–1144.

MacLeod, W. Bentley and James Malcomson, "Motivation and Markets," *American Economic Review*, 1998, *88* (3), 388–411.

Malcomson, James, "Individual Employment Contracts," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3B, North Holland, 1999, pp. 2291–2372.

Manning, Alan, "Imperfect Competition in the Labor Market," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 4B, North Holland, 2011, pp. 973–1041.

Manso, Gustavo, "Motivating Innovation," *The Journal of Finance*, 2011, *66* (5), 1823–1860.

Milgrom, Paul and Sharon Oster, "Job Discrimination, Market Forces, and the Invisibility Hypothesis," *Quarterly Journal of Economics*, 1987, *102* (3), 453–476.

Mirrlees, James, "The Optimal Structure of Incentives and Authority within an Organization," *The Bell Journal of Economics*, 1976, *7* (1), 105–131.

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.

Pissarides, Christopher, *Equilibrium Unemployment Theory*, 2nd ed., Cambridge, Mass: The MIT Press, 2000.

_ , "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?," *Econometrica*, 2009, *77* (5), 1339–1369.

Postel-Vinay, Fabien and Jean-Marc Robin, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

Putnam, Jonathan, "The Value of International Patent Rights," 1996. Yale PhD dissertation.

Robinson, Joan, *The Economics of Imperfect Competition*, London: Macmillan and Company, 1933.

Rosen, Sherwin, "The Theory of Equalizing Differences," in Orley Ashenfelter and Richard Layard, eds., *Handbook of Labor Economics*, Vol. 1, North Holland, 1986, pp. 641–692.

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.

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick, "Capitalists in the Twenty-First Century," 2017. Working paper.

Song, Jae, David Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter, "Firming Up Inequality," 2016. NBER working paper no. 21199.

Sorkin, Isaac, "Ranking Firms Using Revealed Preference," 2016. Working paper.

Stole, Lars and Jeffrey Zwiebel, "Intra-Firm Bargaining under Non-Binding Contracts," *The Review of Economic Studies*, 1996, *63* (3), 375–410.

Suárez Serrato, Juan Carlos and Owen Zidar, "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms," *American Economic Review*, 2016, *106* (9), 2582–2624.

Toivanen, Otto and Lotta Väänänen, "Returns to Inventors," *Review of Economics and Statistics*, 2012, *94* (4), 1173–1190.

Trajtenberg, Manuel, *Economic Analysis of Product Innovation: The Case of CT Scanners*, Cambridge, MA: 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.

Weyl, Glen and Michal Fabinger, "Pass-Through as an Economic Tool: Principles of Incidence under Imperfect Competition," *Journal of Political Economy*, 2013, *121* (3), 528–583.

Williams, Heidi, "How Do Patents Affect Research Investments?," *Annual Review of Economics*, 2017, *9*, 441–469.

Wooldridge, Jeffrey, *Econometric Analysis of Cross Section and Panel Data*, 2nd ed., Cambridge, MA: MIT Press, 2010.

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: KPSS Value (ξ): Predicted Versus Actual

β: 1.03
se: .12

Line of best fit   ----- 45 degree line
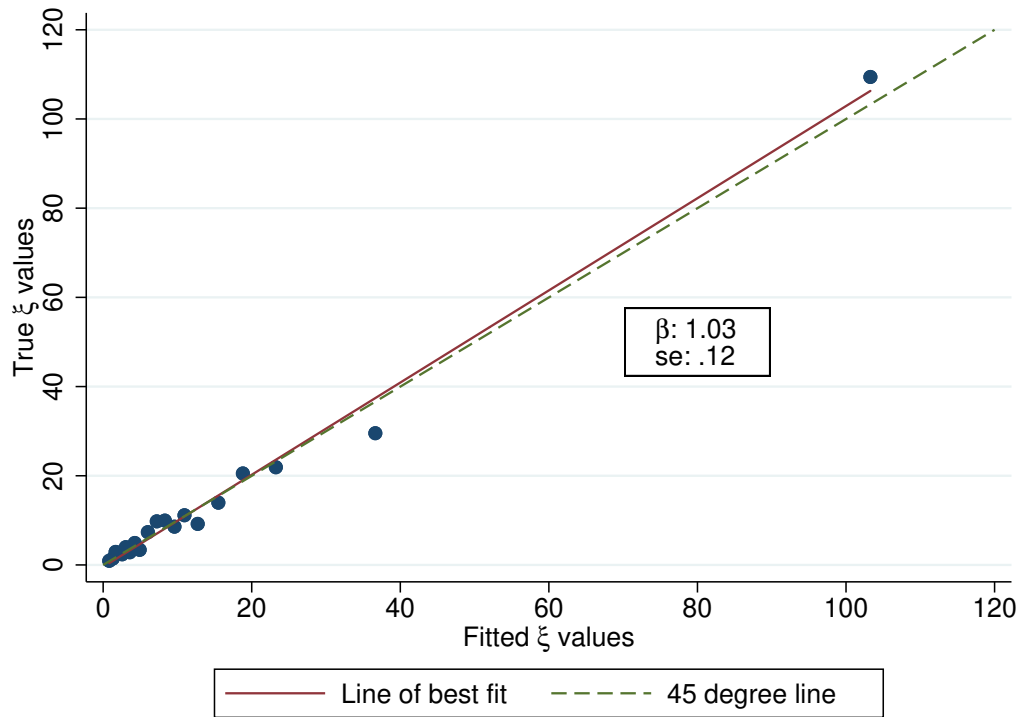
Notes: This figure is a binned scatterplot of actual versus predicted values of the KPSS measure of patent value ξ in millions of 1982 dollars. The sample is the subset of patent applications with non-missing values for the KPSS measure of patent value ξ. Predictions are formed based on estimates from the random effects Poisson model described in Section 5. The data in this figure have been grouped into twenty equal-sized bins. In the microdata, the slope is 1.12, as reported in the text. Here, the coefficient β instead reports the two-stage least squares slope using twenty bin dummies as instruments for predicted values and "se" reports the associated standard error.

Figure 2: Impacts by Predicted Patent Value: Surplus and Wage Bill
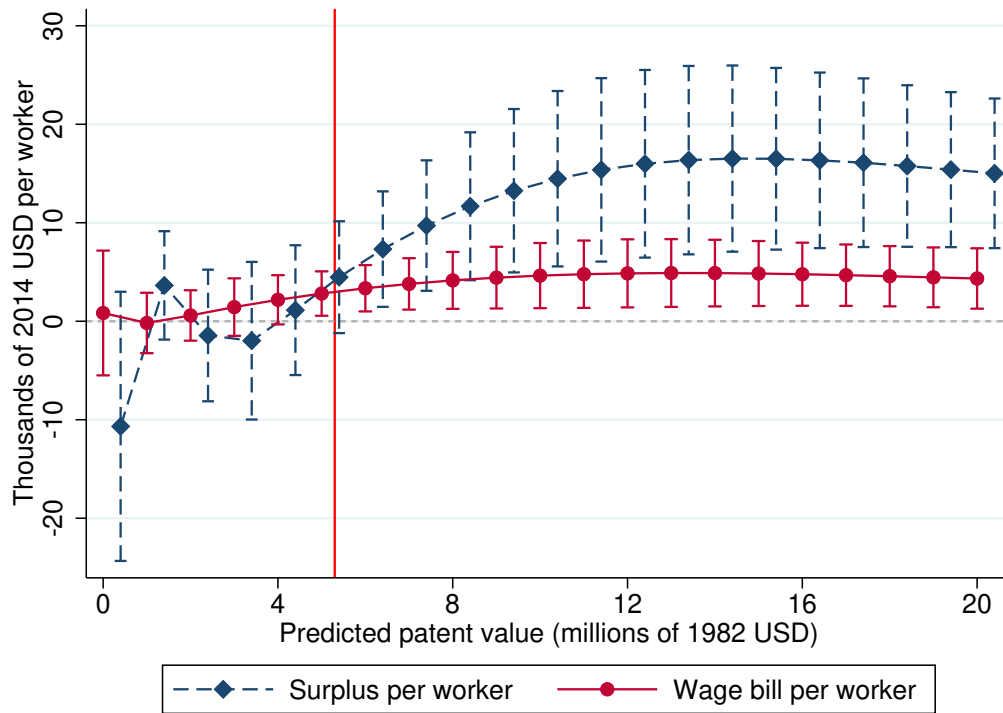
Notes: This figure shows the impact of an initial patent allowance on surplus per worker and wage bill per worker as a function of predicted patent value in our analysis sample. The vertical, red line is the cut-off value for the top-quintile predicted patent value subsample, and is equal to 5.3 million in 1982 USD. Values along the x-axis for the surplus series are offset from their integer value to improve readability. Surplus is EBITD (earnings before interest, tax, and depreciation) + wage bill. 95% confidence intervals shown based upon standard errors two-way clustered by (1) art unit, and (2) application year by decision year.

Figure 3: Event Study Estimates: Surplus (EBITD+Wage Bill) per Worker
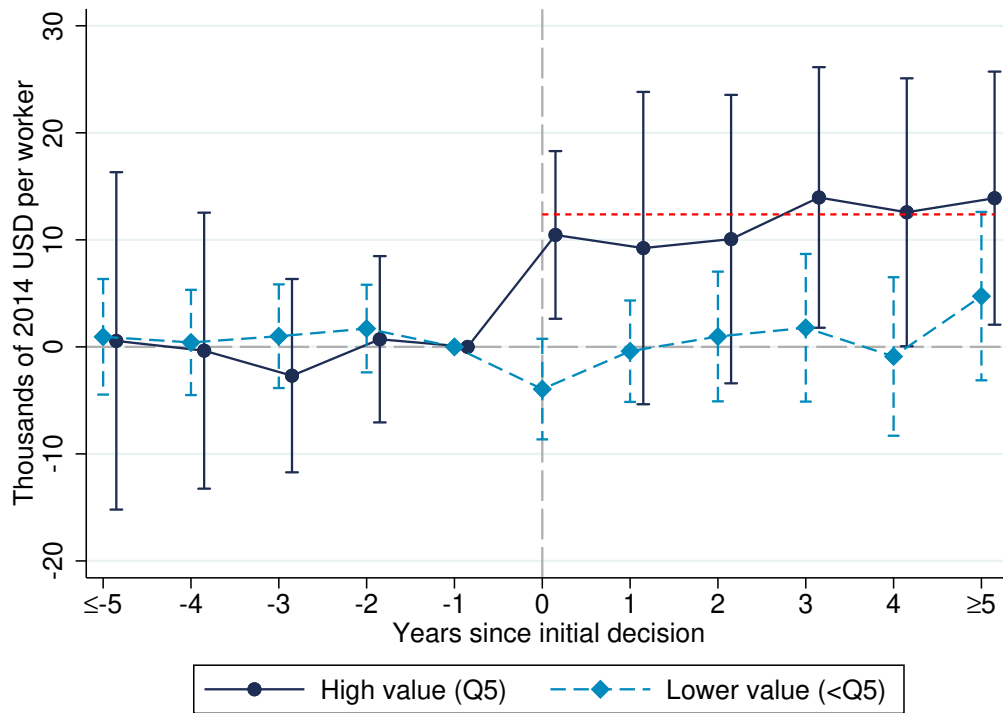
Notes: This figure plots the response of surplus per worker following an initial allowance, separately for high and low ex-ante valuable patent applications, in our analysis sample. Regressions include art unit by application year by calendar year fixed effects and firm fixed effects, as in equation (6). Standard errors are two-way clustered by (1) art unit, and (2) application year by decision year. The horizontal short-dashed, red line is the pooled differences-in-differences estimate of the impact of winning a valuable patent on surplus per worker from Table 5. Surplus is EBITD (earnings before interest, tax, and depreciation) + wage bill. $Q5$ is quintile five of predicted patent value. $< Q5$ are the remaining four quintiles. 95% confidence intervals shown. $Q5$ coefficients are offset from their integer x-axis value to improve readability.

Figure 4: Event Study Estimates: Wage Bill per Worker
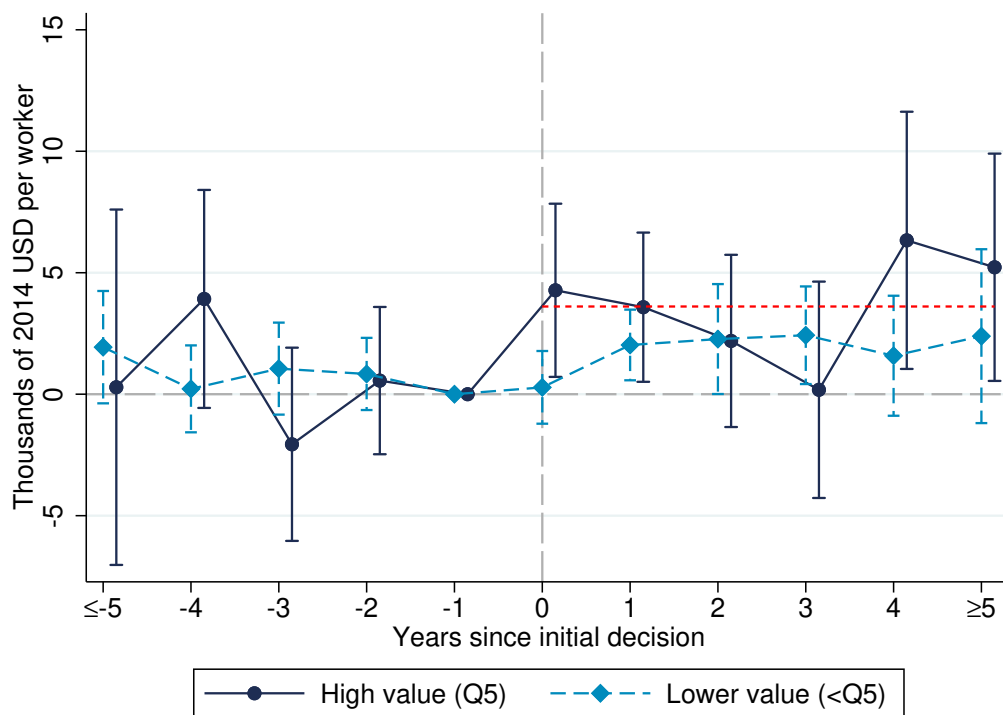
Notes: This figure plots the response of wage bill per worker following an initial allowance, separately for high and low ex-ante valuable patent applications, in our analysis sample. Regressions include art unit by application year by calendar year fixed effects and firm fixed effects, as in equation (6). Standard errors are two-way clustered by (1) art unit, and (2) application year by decision year. The horizontal short-dashed, red line is the pooled differences-in-differences estimate of the impact of winning a valuable patent on wage bill per worker from Table 5. Q5 is quintile five of predicted patent value. < Q5 are the remaining four quintiles. Q5 coefficients are offset from their integer x-axis value to improve readability. 95% confidence intervals shown.

Figure 5: Event Study Estimates: log(employees)

Notes: This figure plots the response of the logarithm of employees per worker following an initial allowance, separately for high and low ex-ante valuable patent applications, in our analysis sample. Regressions include art unit by application year by calendar year fixed effects and firm fixed effects, as in equation (6). Standard errors are two-way clustered by (1) art unit, and (2) application year by decision year. The horizontal short-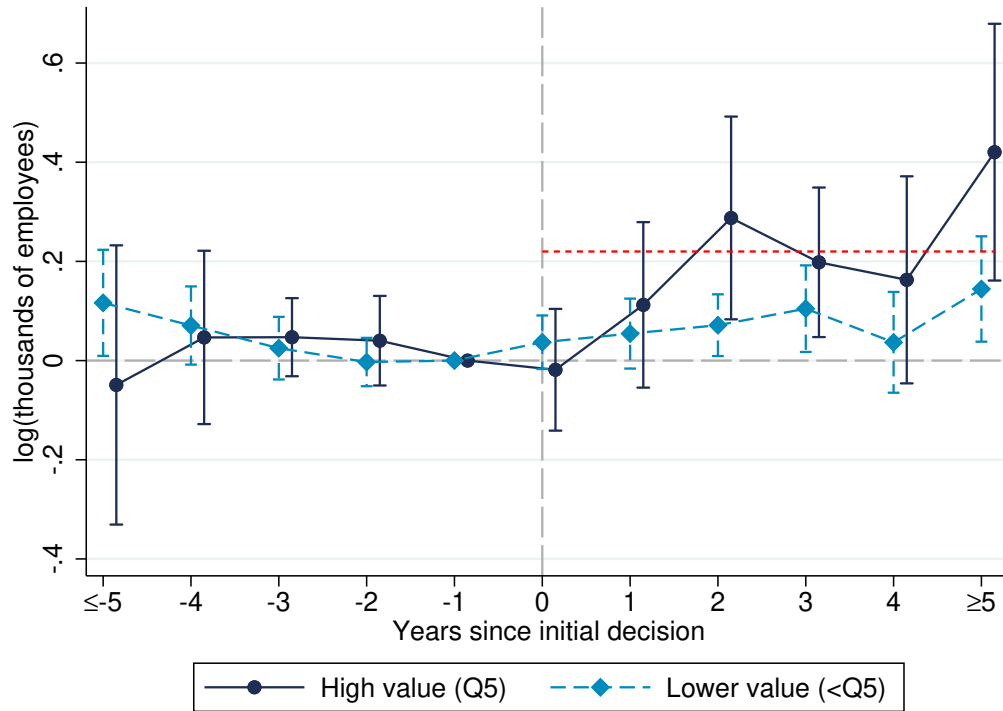dashed, red line is the pooled differences-in-differences estimate of the impact of winning a valuable patent on the logarithm of the number of employees at the firm in thousands of people. $Q5$ is quintile five of predicted patent value. $< Q5$ are the remaining four quintiles. 95% confidence intervals shown. Values along the x-axis for the difference in $Q5$ are offset from their integer value to improve readability.

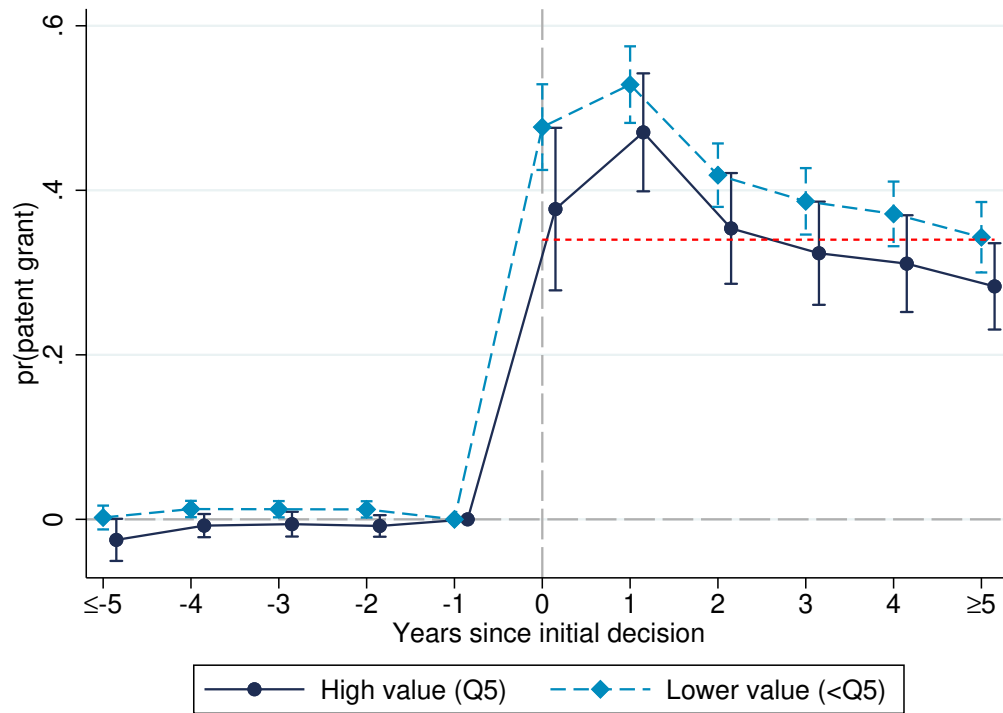Figure 6: Event Study Estimates: Probability of Patent Grant

Notes: This figure plots the response of the probability of patent grant following an initial allowance, separately for high and low ex-ante valuable patent applications, in our analysis sample. Regressions include art unit by application year by calendar year fixed effects and firm fixed effects, as in equation (6). Standard errors are two-way clustered by (1) art unit, and (2) application year by decision year. The horizontal short-dashed, red line is the pooled difference-in-differences estimate of the impact of winning a valuable patent on the probability of the patent having been granted. Values along the x-axis for the difference in $Q5$ are offset from their integer value to improve readability. $Q5$ is quintile five of predicted patent value. $< Q5$ are the remaining four quintiles. 95% confidence intervals shown.

Figure 7: Earnings by Quartile: Levels and sinh$^{-1}(\cdot)$

(a) Wage Bill and Stayer Earnings: Levels

(b) Entrant Earnings: Levels

(c) Wage Bill and Stayer Earnings: sinh$^{-1}(\cdot)$

(d) Entrant Earnings: sinh$^{-1}(\cdot)$
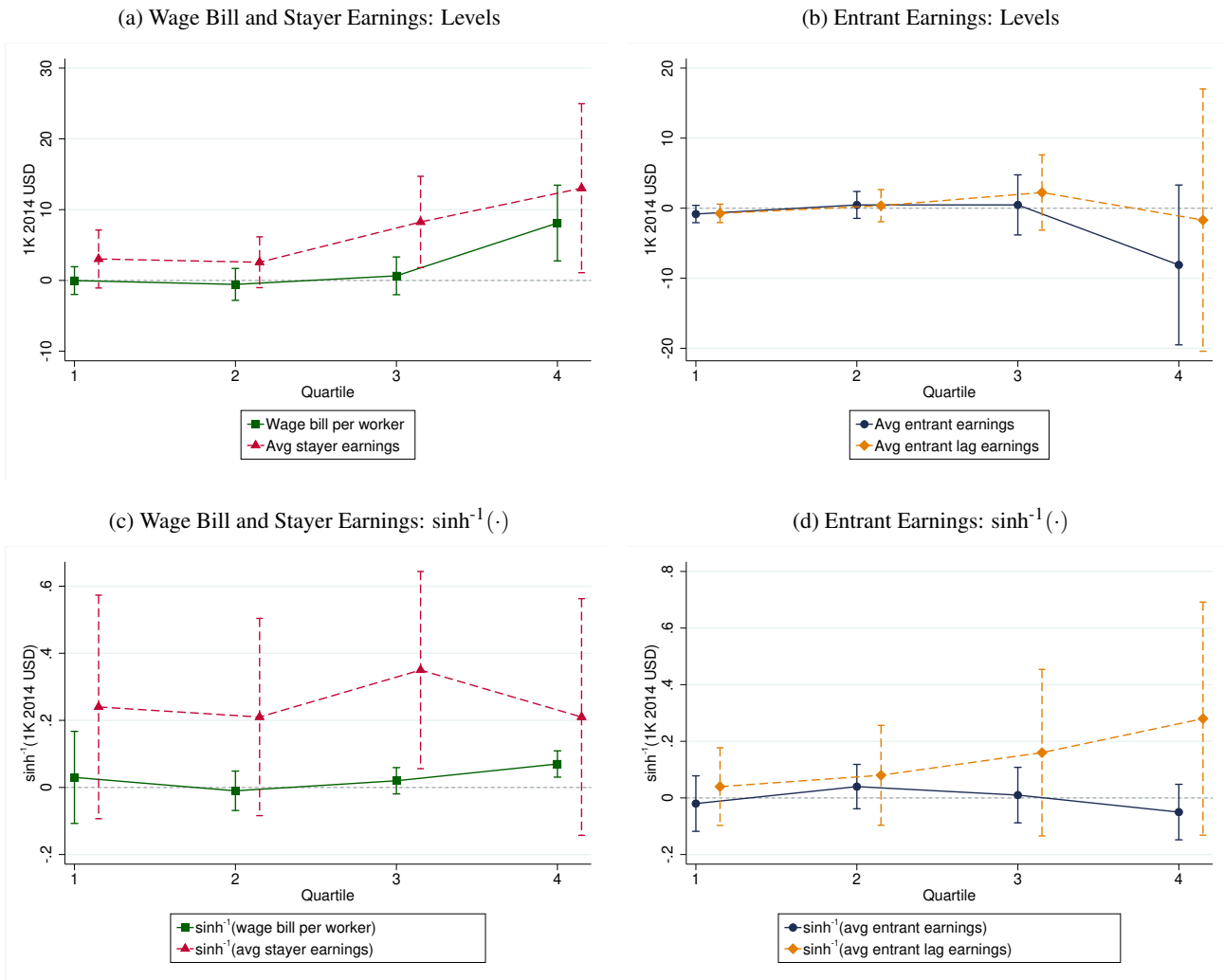
Notes: This figure shows the differences-in-differences estimates of the effect of valuable initial allowances on different measures of employee earnings by within-firm wage quartiles. Estimates correspond to coefficients on interactions of an indicator for high-value patent applications with a post-decision indicator and an indicator for the application being initially allowed. Each series is from a separate pooled difference-in-differences regression using firm wage quartile specific outcomes and regressors, as in equation (7). Controls include main effect of value category interacted with a post-decision indicator, firm fixed effects, and Art Unit by application year by calendar year fixed effects. Two-way standard errors are clustered by (1) art unit, and (2) application year by decision year. Quartiles along the x-axis in the second series of each panel are offset from their integer value to improve readability. Earnings and wage bill are measured in thousands of 2014 USD. Stayers are defined as those who were employed by the same firm in the year of application. Panels (a) and (c) report impacts of winning a valuable patent allowance for wage bill and stayer earnings in levels and for sinh$^{-1}(\cdot)$ transformations, respectively, by within-firm wage quartiles. Panels (b) and (d) report impacts of winning a valuable patent allowance for entrant earnings and entrant earnings in the year before entry in levels and for sinh$^{-1}(\cdot)$ transformations, respectively, by within-firm wage quartiles. Each panel is from a different subsample of the analysis sample that retains firms with at least four employees. Stayers are defined as those who were employed by the same firm in the year of application. Entrants are defined as those employees who were not employed at the firm in the previous year. Stayer data series in panels (a) and (c) additionally require at least one stayer in the relevant application year quartile. Entrant data series in panels (b) and (d) additionally require at least one entrant in the relevant earnings year quartile.

Table 1: Sample Construction

	Application-assignee pairs	Applications	Assignees	EINs
Panel A: USPTO sample				
Full sample	3,737,351	3,601,913	317,370	—
Filed between 2000 and 2010	3,063,980	2,954,507	279,936	—
Non-missing assignees	2,708,829	2,599,373	279,935	—
Non-child applications	1,341,843	1,295,649	130,619	—
Utility applications	1,339,146	1,293,054	130,113	—
First application by assignee	130,113	125,018	130,113	—
No prior grant to assignee	99,871	95,767	99,871	—
Panel B: USPTO-tax merge	—	39,452	39,814	81,934
First application by EIN	—	37,714	—	81,877
No prior grant to EIN	—	35,643	—	78,291
EIN with largest revenue	—	35,643	—	35,643
Active firms	—	9,735	—	9,735

Notes: This table describes the construction of our analysis sample. When selecting the first application by each assignee by date of filing ("First application by assignee"), ties are broken by taking the smallest application number. When selecting the first application for each EIN ("First application by EIN"), we drop EINs with more than one first application. When removing assignees ("No prior grant to assignee") and EINs ("No prior grant by EIN") with prior grants, we do so by checking against the assignees and EINs for the census of patents granted since 1976 and filed before 29 November 2000. When selecting the EIN with the largest revenue ("EIN with largest revenue"), we compare based on the revenue in the year of the application. Active firms are defined as EINs with non-zero/non-missing total income or total deductions in the application year and in the three previous years, a positive number of employees in the application year, and revenue less than 100 million in 2014 USD.

Table 2: Summary Statistics

	1. Analysis sample				2. Top-quintile dosage sample			
	Mean	p10	p50	p90	Mean	p10	p50	p90
Panel A: firm outcomes								
Revenue	9,840	226.65	3,234	29,082	12,922	347.77	4,855	39,614
Value added	3,950	117.78	1,299	10,456	5,278	170.59	1,948	14,405
EBITD	103.01	-916.47	87.2	1,979	11.45	-1,858.87	110.17	2,741
Employment	45.64	2	17.08	117.17	61.54	2.93	25.33	152.37
Value added per worker	116.36	15.66	84.7	287.36	119.61	13.97	86.45	313.18
EBITD per worker	0.8	-70.88	6.85	69.53	0.44	-76.98	6.94	77.21
Predicted patent value	4,910	216	1,743	10,912	18,066	6,077	10,455	36,441
% Patents initially allowed	8.4	.	.	.	7.7	.	.	.
Panel B: worker outcomes								
Labor compensation	2,131	67.96	703.76	5,410	3,089	106.7	1,113	8,310
Wage bill	2,412	71.4	841.25	6,487	3,375	116.58	1,361	9,425
Labor compensation per worker	57.21	10.61	43.12	136.34	60.71	12.02	47.46	148.37
Wage bill per worker	54.99	17.94	47.95	109.33	58.38	22.31	51.52	114.95
% Female employment	30.2	0	25.6	66.1	29.8	0	26.3	62.6
% Contractors	18.4	0	10.5	53	17.8	0	11.1	50.3
% Entrants	28.8	0	24.6	62.9	30	0	25.4	63
% Inventors	11.4	0	2.5	36.4	10.5	0	3.3	32.4
firm observations	9,735				1,947			

Notes: This table tabulates summary statistics on firm (Panel A) and worker (Panel B) outcomes for each of two samples: our analysis sample of matched patent application-firm pairs, and our sub-sample of patent application-firm pairs in the analysis sample for which the patent applications are in the top quintile of predicted value ("dosage"). All summary statistics are measured in the year in which the patent application was filed. For each variable in each panel for each sample, we tabulate the mean, 10th percentile, 50th percentile (median), and 90th percentile; to protect taxpayer anonymity, p10, p50, and p90 refer to centile means. Entrants are defined as those employees who were not employed at the firm in the previous year. EBITD is earnings before interest, taxes, and deductions. Revenue, value added, EBITD, labor compensation, and wage bill are reported in thousands of 2014 USD. Predicted patent value is reported in thousands of 1982 USD. All percent or "per worker" outcomes in Panel B are per W2 employee except for "% Contractors" which is per W2 employee + 1099 independent contractor. For all other variable definitions, please see the text of the paper and Appendix A.

44

Table 3: Balance of Assignee Characteristics Across Initially Allowed and Initially Rejected Patent Applications

| | Initially allowed | | | |
| | Analysis sample | | Top quintile | |
	(1a)	(1b)	(2a)	(2b)
log(employees)	-3.70	-2.06	-0.19	1.78
	(1.86)	(2.18)	(4.74)	(4.77)
Revenue per worker	0.03	0.02	0.01	0.01
	(0.01)	(0.02)	(0.03)	(0.03)
Value added per worker	-0.14	-0.07	-0.07	0.00
	(0.05)	(0.06)	(0.08)	(0.12)
Wage bill per worker	0.14	0.14	0.08	-0.11
	(0.10)	(0.13)	(0.21)	(0.24)
EBITD per worker	0.11	0.06	0.18	0.12
	(0.05)	(0.07)	(0.10)	(0.14)
Observations	9,735	8,649	1,947	1,665
AU-AY FEs		✓		✓
p-value	0.004	0.481	0.515	0.808

Notes: This table reports covariate balance tests for initial patent allowances. Specifically, the coefficients report linear probability model estimates of the marginal effect of the included covariate on the probability that a patent application receives an initial allowance; all coefficients have been multiplied by 1,000 for ease of interpretation. AU-AY FEs denotes the inclusion of Art Unit (AU) by application year (AY) fixed effects. Covariates are measured as of the year of application. Columns (1a) and (1b) report the results for observations in the analysis sample. Columns (2a) and (2b) report the results for observations in the top-quintile predicted patent value. Singleton observations are dropped in the fixed effects specifications, which accounts for the smaller number of observations in Column (1b) relative to Column (1a) and in Column (2b) relative to Column (2a). Standard errors (reported in parentheses) are two-way clustered by art unit and application year by decision year except in column (2b) which clusters by art unit (because the estimated two-way variance covariance matrix was singular). The p-value reports the probability that the covariates measured in the year of application do not influence the probability of an initial allowance. EBITD is earnings before interest, taxes, and deductions. Revenue, value added, wage bill, and EBITD are measured in thousands of 2014 USD.

Table 4: Prediction of KPSS Patent Value Based on Patent Application and Assignee Characteristics

	KPSS value (ξ)	
$\mathbf{1}$(patent family size $= 1$)	0.27	(0.06)
log(patent family size)	0.22	(0.04)
$\mathbf{1}$(number of claims $= 1$)	0.69	(0.19)
log(number of claims)	0.31	(0.03)
$\mathbf{1}$(revenue $= 0$)	1.41	(0.14)
log(revenue)	0.14	(0.02)
$\mathbf{1}$(employees $= 0$)	0.45	(0.07)
log(employees)	0.00	(0.02)
Application year	-0.02	(0.05)
(Application year)^2	-0.01	(0.01)
Decision year	0.31	(0.06)
(Decision year)^2	-0.03	(0.01)
Constant	-1.40	(0.21)
log(σ)	0.24	(0.05)
Observations	596	
Art units	260	
χ^2	10,362	

Notes: This table reports the relationship between KPSS ξ patent value, and patent application and firm level co-variates. Coefficient estimates are from a Poisson model with art unit random effects. The sample is the subsample of granted patents for which the Kogan et al. (2017) measure of patent value is available in our analysis sample, except we retain firms with more than 100 million in 2014 revenue (unlike in our analysis sample) in order to maximize sample size (N=596). The dependent variable is the KPSS measure of patent value ξ in millions of 1982 dollars. Standard errors are reported in parentheses. Patent family size measures the number of countries in which the patent application was submitted. Number of claims measures the number of claims in the published US patent application. Revenue (in thousands of 2014 dollars) and number of employees are measured as of the year the US patent application was filed. log(σ) reports the log of the estimated variance of the art unit random effects. χ^2 reports the results of a likelihood ratio test statistic against a restricted Poisson model without art unit random effects; this test has one degree of freedom.

46

Table 5: Impacts on Firm Aggregates

	(1) Positive employ-ment	(2) Firm size	(3) Revenue per worker	(4) Value added per worker	(5) EBITD per worker	(6) Wage bill per worker	(7) Surplus per worker	(8) Labor comp per worker	(9) W2 + 1099 per worker	(10) Income tax per worker
Panel A: levels										
High value ($Q5$)	0.00	18.79	37.25	15.90	9.11	3.61	12.38	4.00	2.82	0.84
	(0.04)	(9.36)	(14.71)	(5.11)	(3.85)	(1.51)	(3.55)	(2.67)	(1.49)	(0.65)
Lower value ($<Q5$)	-0.01	-2.15	-9.72	0.87	-1.37	0.78	-0.24	1.24	0.49	0.73
	(0.01)	(2.95)	(8.38)	(3.84)	(1.79)	(0.89)	(2.08)	(1.65)	(0.85)	(0.56)
Panel B: sinh$(\cdot)^{-1}$										
High value ($Q5$)	.	0.22	0.04	0.10	0.34	0.08	0.44	0.03	0.07	0.08
	.	(0.08)	(0.14)	(0.11)	(0.22)	(0.02)	(0.15)	(0.10)	(0.04)	(0.02)
Lower value ($<Q5$)	.	0.03	-0.01	0.02	-0.07	0.00	-0.09	0.01	-0.01	0.05
	.	(0.04)	(0.05)	(0.06)	(0.11)	(0.02)	(0.09)	(0.05)	(0.02)	(0.03)
Observations	155,682	103,459	103,459	103,459	103,459	103,459	103,459	103,459	107,811	103,181

Notes: This table reports difference–in–differences estimates of the effect of initial allowances on firm and worker outcomes, separately for high and low ex-ante valuable patent applications, in our analysis sample. Estimates correspond to coefficients on interactions of the designated value category with a post-decision indicator and an indicator for the application being initially allowed. Controls include main effect of value category interacted with a post-decision indicator, firm fixed effects, and Art Unit by application year by calendar year fixed effects, as in equation (7). Standard errors (reported in parentheses) are two-way clustered by (1) art unit, and (2) application year by decision year. Panel B documents impacts on sinh$^{-1}(\cdot)$ transformations of the dependent variables (see Appendix D). EBITD is earnings before interest, taxes, and deductions. Surplus is EBITD + wage bill. Labor compensation measures total deductions for labor expenses claimed by the firm. "W2 + 1099" measures to the sum of W2 and 1099 earnings divided by the sum of the number of W2's and 1099's filed. "Income tax per worker" is the average worker's individual income tax liability. Revenue, value added, EBITD, wage bill, surplus, labor compensation, and W2 + 1099 pay are reported in thousands of 2014 USD.

Table 6: Workforce Composition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Share female	Share inventors	Avg entrant earnings (yr bef ent)	Avg separator earnings (yr bef sep)	Stayers	Share stayers	Avg stayer earnings (in app yr)	Avg age
Panel A: levels								
High value (Q5)	-0.01	-0.01	-0.90	0.61	0.69	-0.02	1.29	-1.08
	(0.01)	(0.01)	(2.03)	(1.15)	(3.85)	(0.02)	(1.60)	(0.55)
Lower value (< Q5)	-0.01	-0.01	0.45	-0.02	0.07	0.00	0.97	0.07
	(0.01)	(0.01)	(0.68)	(0.52)	(1.79)	(0.01)	(1.19)	(0.22)
Panel B: sinh(·)$^{-1}$								
High value (Q5)	-0.01	-0.01	0.06	0.03	0.12	-0.02	0.02	-0.02
	(0.01)	(0.01)	(0.10)	(0.05)	(0.05)	(0.02)	(0.03)	(0.01)
Lower value (< Q5)	-0.01	-0.01	-0.01	-0.01	0.01	0.00	0.01	0.00
	(0.01)	(0.01)	(0.04)	(0.02)	(0.03)	(0.01)	(0.02)	(0.00)
Observations	103,459	103,459	70,086	75,541	103,459	103,459	99,580	103,456

Notes: This table reports difference-in-differences estimates of the effect of initial allowances on within-firm workforce composition measures, separately for high and low ex-ante valuable patent applications, in our analysis sample. Estimates correspond to coefficients on interactions of the designated value category with a post-decision indicator and an indicator for the application initially allowed. Controls include main effect of value category interacted with a post-decision indicator, firm fixed effects, and Art Unit by application year by calendar year fixed effects, as in equation (7). Standard errors (reported in parentheses) are two-way clustered by (1) art unit, and (2) application year by decision year. Panel B documents impacts on sinh^{-1}(·) transformations of the dependent variables (see Appendix D). "Avg entrant earnings (yr bef ent)" measures the earnings of entrants in the year before they joined the firm. "Avg separator earnings (yr bef sep)" measures the earnings of separators in the year before they leave the firm. Stayers are defined as those who were employed by the same firm in the year of application. Entrants are defined as those employees who were not employed at the firm in the previous year. "Avg age" measures the average age of all employees at the firm. Earnings are measured in thousands of 2014 USD.

Table 7: Within-Firm Inequality

	(1) Avg male earnings	(2) Avg female earnings	(3) Gender earnings gap	(4) Avg inventor earnings	(5) Avg earnings of non-inventors	(6) Inventor earnings gap	(7) Avg non-inventor male earnings	(8) Wage bill per worker (Q_1)	(9) Wage bill per worker (Q_4)	(10) Wage bill per worker ($Q_4 - Q_1$)
Panel A: levels										
High value ($Q5$)	5.81	0.17	6.98	16.61	2.20	14.64	3.55	-0.02	8.11	8.14
	(1.94)	(1.43)	(1.98)	(8.48)	(1.39)	(7.73)	(1.66)	(1.00)	(2.73)	(2.52)
Lower value ($< Q5$)	0.33	-0.47	-0.08	-1.18	0.46	-1.69	-0.38	0.06	2.77	2.72
	(1.15)	(0.50)	(1.06)	(4.58)	(0.80)	(4.68)	(0.96)	(0.32)	(2.40)	(2.34)
Panel B: $\sinh(\cdot)^{-1}$										
High value ($Q5$)	0.09	0.01	0.10	0.10	0.06	0.12	0.06	0.03	0.07	0.04
	(0.03)	(0.04)	(0.05)	(0.06)	(0.03)	(0.06)	(0.03)	(0.07)	(0.02)	(0.07)
Lower value ($< Q5$)	0.00	-0.03	0.01	0.00	-0.01	-0.01	-0.02	0.00	0.01	0.01
	(0.02)	(0.02)	(0.02)	(0.04)	(0.02)	(0.04)	(0.02)	(0.03)	(0.02)	(0.03)
Observations	95,026	84,578	80,238	52,471	100,923	50,045	90,617	82,763	81,549	81,549

Notes: This table reports difference-in-differences estimates of the effect of initial allowances on within-firm inequality measures, separately for high and low ex-ante valuable patent applications, in our analysis sample. Estimates correspond to coefficients on interactions of the designated value category with a post-decision indicator and an indicator for the application being initially allowed. Controls include main effect of value category interacted with a post-decision indicator, firm fixed effects, and Art Unit by application year by calendar year fixed effects, as in equation (7). Standard errors (reported in parentheses) are two-way clustered by (1) art unit, and (2) application year by decision year. Panel B documents impacts on $\sinh^{-1}(\cdot)$ transformations of the dependent variables, except in Columns 3 and 6 where the outcome variable is the impact on the *difference* in the $\sinh^{-1}(\cdot)$ transformation of group means. "Gender earnings gap" measures the difference between average male and female earnings at firms where both genders are present. "Inventor earnings gap" measures the difference between average inventor and non-inventor earnings at firms where both inventors and non-inventors are present. "Wage bill per worker (Q_1)" measures the average wage bill in within-firm wage quartile one. "Wage bill per worker (Q_4)" measures the average wage bill in within-firm wage quartile four. "Wage bill per worker ($Q_4 - Q_1$)" measures the difference between average Q_4 and Q_1 earnings. Entrants are defined as those employed by the same firm in the year of application. Stayers are defined as those who were employed by the same firm in the previous year. Earnings and wage bill are measured in thousands of 2014 USD.

49

Table 8: Earnings Impacts by Year of Entry/Exit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
							Change since application year		
	Avg cohort earnings	Avg stayer earnings	Avg leaver earnings	Avg entrant earnings	Avg recent entrant earnings	Avg post-app hire earnings	Avg stayer earnings	Avg leaver earnings	Avg entrant earnings
Panel A: levels									
High value ($Q5$)	3.91	7.78	-1.52	0.11	-2.67	-4.03	6.48	2.90	1.01
	(2.31)	(2.93)	(1.94)	(1.62)	(1.83)	(2.10)	(3.14)	(5.36)	(1.80)
Lower value ($<Q5$)	0.35	2.46	0.90	0.28	0.79	0.30	1.49	-3.88	-0.16
	(1.18)	(1.58)	(1.40)	(0.78)	(1.00)	(0.95)	(1.64)	(2.44)	(0.70)
Panel B: sinh$(\cdot)^{-1}$									
High value ($Q5$)	0.06	0.10	-0.04	0.03	-0.03	-0.06	0.08	0.08	-0.03
	(0.03)	(0.04)	(0.05)	(0.05)	(0.04)	(0.06)	(0.04)	(0.07)	(0.08)
Lower value ($<Q5$)	0.00	0.02	0.00	0.01	0.01	0.01	0.01	-0.05	0.02
	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.04)	(0.04)
Observations	151,928	99,580	109,196	70,086	68,698	47,323	99,580	109,196	70,086

Notes: This table reports difference-in-differences estimates of the effect of initial allowances on worker outcomes for employees who stay, enter, and exit $\hat{\xi}$, separately for high and low ex-ante valuable patent applications. Estimates correspond to coefficients on interactions of the designated value category with a post-decision indicator and an indicator for the application being initially allowed. Controls include main effect of value category interacted with a post-decision indicator, firm fixed effects, and Art Unit by application year by calendar year fixed effects, as in equation (7). Standard errors (reported in parentheses) are two-way clustered by (1) art unit, and (2) application year by decision year. Panel B documents impacts on sinh$^{-1}(\cdot)$ transformations of the dependent variables (see Appendix D). "Avg cohort earnings" measures the W2 earnings of workers employed by the firm in the year of application. "Avg stayer earnings" measures the W2 earnings of workers employed by the firm in the year of application who are also employed in the present year. "Avg leaver earnings" measures the W2 earnings of workers employed by the firm in the year of application who are not employed in the present year. "Avg entrant earnings" measures the W2 earnings of employees who were not employed by the firm in the previous year. "Avg recent entrant earnings" measures the W2 earnings of employees who were hired by the firm within the past three years. "Avg post-app hire earnings" tracks the earnings of employees hired by the firm following the application year. "Change since application year" columns are earnings in the current year (Columns 7 and 8), and year of entry (Column 9) minus the respective values in the application year. Stayers are defined as those who were employed by the same firm in the year of application. Entrants are defined as those employees who were not employed at the firm in the previous year. Earnings are measured in thousands of 2014 USD.

50

Table 9: Pass-Through Estimates

	Wage bill per worker		Avg stayer earnings		Avg earnings of stayers minus earnings in appl yr		Avg non-inventor stayer earnings	
	(1a) OLS	(1b) IV	(2a) OLS	(2b) IV	(3a) OLS	(3b) IV	(4a) OLS	(4b) IV
Panel A: levels								
Surplus per worker	0.16	0.29	0.20	0.60	0.19	0.50	0.16	0.49
	(0.01)	(0.12)	(0.01)	(0.30)	(0.01)	(0.27)	(0.00)	(0.22)
1st stage F		12.18		13.46		13.46		8.94
Exogeneity		0.293		0.137		0.218		0.074
Panel B: sinh$(\cdot)^{-1}$								
Surplus per worker	0.04	0.19	0.04	0.23	0.04	0.18	0.04	0.20
	(0.00)	(0.09)	(0.00)	(0.11)	(0.00)	(0.10)	(0.00)	(0.11)
1st stage F		9.02		10.63		10.63		9.05
Exogeneity		0.055		0.054		0.108		0.064
observations	103,459	103,459	99,580	99,580	99,580	99,580	94,931	94,931

Notes: This table reports OLS and IV estimates of the effect of increases in surplus per worker on selected earnings outcomes. The excluded instrument is the interaction of top quintile of ex-ante value $\hat{\xi}$ category with a post-decision indicator and an indicator for the application being initially allowed. Controls include main effect of value category interacted with post-decision indicator and interaction of lower quintile value category with a post-decision indicator interacted with an indicator for initially allowed, firm fixed effects, and Art Unit by application year by calendar year fixed effects. Standard errors (reported in parentheses) are two-way clustered by (1) art unit, and (2) application year by decision year. Panel B takes sinh$^{-1}(\cdot)$ transformations of the outcome and endogenous variables (see Appendix D). "Exogeneity" reports the p-value from a test of the null hypothesis that IV and OLS estimators have same probability limit. Stayers are defined as those who were employed by the same firm in the year of application. Surplus is EBITD (earnings before interest, tax, and depreciation) + wage bill. Earnings, wage bill, and surplus are measured in thousands of 2014 USD.

Table 10: Labor Supply

	sinh^{-1}(employees)		sinh^{-1}(males)		sinh^{-1}(non-inventors)		sinh^{-1}(stayers)			
	(1a) OLS	(1b) IV	(2a) OLS	(2b) IV	(3a) OLS	(3b) IV	(4a) OLS	(4b) IV	(5a) OLS	(5b) IV
sinh^{-1}(wage bill per worker)	0.19 (0.02)	2.70 (1.41)								
sinh^{-1}(avg male earnings)			0.10 (0.02)	1.76 (1.38)						
sinh^{-1}(avg non-inventor earnings)					0.20 (0.03)	3.49 (2.22)				
sinh^{-1}(avg stayer earnings)							0.12 (0.02)	1.16 (0.57)		
sinh^{-1}(avg stayer earnings) - sinh^{-1}(avg stayer pre earnings)									0.33 (0.02)	1.49 (0.78)
Observations	103,459	103,459	95,026	95,026	100,923	100,923	99,580	99,580	99,580	99,580
1$^{\text{st}}$ stage F		10.80		7.93		4.95		8.10		4.79
Exogeneity		0.043		0.116		0.044		0.040		0.065

Notes: This table reports OLS and IV estimates of the effect of increases in selected earnings measures on selected measures of employment. The excluded instrument is the interaction of top quintile of ex-ante value $\hat{\xi}$ category with a post-decision indicator and an indicator for the application being initially allowed. Controls include main effect of value category interacted with post-decision indicator and interaction of lower quintile value category with a post-decision indicator interacted with an indicator for initially allowed, firm fixed effects, and Art Unit by application year by calendar year fixed effects. Standard errors (reported in parentheses) are two-way clustered by (1) art unit, and (2) application year by decision year. The sinh^{-1}($\cdot$) transformation is applied to each of the outcome and endogenous variables (see Appendix D). "Exogeneity" reports p-value for test of null hypothesis that IV and OLS estimators have same probability limit. Stayers are defined as those who were employed by the same firm in the year of application. Earnings and wage bill are measured in thousands of 2014 USD.

52

A   Appendix: Data

A.1   Description of patent data

Our patent data build draws on several sources. Three identification numbers are relevant when using 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 associated with more than one granted patent.

Traditionally, unsuccessful patent applications were not published by the USPTO. However, as part of the American Inventors Protection Act of 1999, the vast majority of patent applications filed in the US on or after 29 November 2000 are published eighteen months after the filing date. There are two exceptions. First, applications granted or abandoned before eighteen months do not appear in this sample unless the applicant chooses to ask for early publication. Lemley and Sampat (2008) estimate that about 17 percent of patents are granted before eighteen months, of which about half (46 percent) are published pre-patent grant. Second, applications pending more than eighteen months can "opt out" of publication if they do not have corresponding foreign applications, or if they have corresponding foreign applications but also have priority dates predating the effective date of the law requiring publication (Lemley, and Sampat 2008).[21]

1. <u>Census of published USPTO patent applications.</u> We observe the census of published (accepted and rejected) patent applications published by the US Patent and Trademark Office (USPTO). Our source for this data is a set of bulk XML files hosted by Google.[22] 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 other USPTO administrative data wherever possible. These records are at the publication number level.

2. <u>Census of granted USPTO patents.</u> For the published USPTO patent applications in our data, we wish to observe which of those applications were granted patents. Our source for this data is a set of bulk XML files hosted by Google.[23] 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 other USPTO administrative data wherever possible. As one specific example, even though patent numbers uniquely identify patent grants, there are twenty-one patent numbers in this data that appear in the data twice with different grant dates. Checking these patent numbers on the USPTO's online Patent Full Text (PatFT) database reveals that in each of these cases, the duplicated patent number with the earlier grant date is correct.[24] Accordingly, we drop the twenty-one observations with the later grant dates.

3. <u>USPTO patent assignment records.</u> Some of our published patent applications are missing assignee information. (Applicants are not required to submit assignee information to the USPTO at the time of application.) Based on informal conversations with individuals at the USPTO, we fill in missing assignee names to the extent possible using the USPTO Patent Assignment data. The USPTO Patent Assignment data records assignment transactions, which are legal transfers of all or part of the right, title, and interest in a patent or application from one or more existing owner to one or more recipient. The dataset is hosted on the USPTO

---

[21]For more details, see `http://www.uspto.gov/web/offices/pac/mpep/s1120.html` and the discussion in Lemley and Sampat (2010). Most applications not published eighteen months after filing are instead published sixty months after filing.

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

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

[24]PatFT can be accessed at `http://patft.uspto.gov/`.

website.[25] Each transaction is associated with a patent number, application number, and/or publication number (wherever each is applicable). The patent assignment records include both initial assignments and re-assignments, but only the former is conceptually appropriate for our analysis since we want to measure invention ownership at the time of application. We isolate initial assignments by taking the assignment from this database with the earliest execution date. If a given assignment has more than one execution date (e.g., if the patent application is assigned to more than one entity), we use the latest execution date within that assignment as the transaction execution date. Using these initial assignments, we fill in assignee organization name as well as assignee address information where possible when these variables are missing from our published patent applications data.

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 datasets we analyze. 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.[26] Our versions of these files were downloaded on 24 March 2014 and are up to date as of February 2014. These records are at the publication number level.

5. USPTO PAIR records. We analyze several variables, such as the date of initial decisions, from the USPTO Patent Application Information Retrieval (PAIR) data, which we draw from an administrative dataset called the Patent Examination Research Dataset (PatEx).[27] With the exception of 264 published patent applications, these data are available for our full sample of published USPTO patent applications. These records are at the application number level.

6. Examiner art unit and pay scale data. Frakes and Wasserman (2017) generously provided us with examiner art unit and GS pay scale data they received through FOIA requests. These data allow us to identify which examiners were active in each art unit in each year.

7. Thomson Innovation database. All of the databases listed above record information obtained directly 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., the 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.[28] 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), 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.

8. 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

---

[25]Available at: https://www.uspto.gov/learning-and-resources/electronic-data-products/patent-assignment-dataset.

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

[27]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.

[28]See http://info.thomsoninnovation.com/.

of that work, the authors constructed technology categories to describe the broad content area of different patents, based on categorizations of the patent technology class and subclass variables.[29] 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 level.

9. <u>Kogan et al. (2017) patent value data</u>. Kogan et al. (2017) provide their final estimates of patent value for their sample of granted patent at `https://iu.app.box.com/v/patents/`. In particular we downloaded the "patents.zip" file, which contains a linkage between USPTO patent grant numbers and the estimate of the patent value ξ. These data were downloaded on—and are accurate as of—7 August 2016. To develop a measure of patent value at the application number level, we associate each application with its potentially numerous patent numbers. We then sum the values of ξ by application number to obtain a measure of the ex-post value of granted applications.

10. <u>USPTO technology center data.</u> Technology centers are groupings of examiner art units. The USPTO hosts a listing of all technology centers and associated examiner art units at `http://www.uspto.gov/patent/contact-patents/patent-technology-centers-management`. We use these groupings to examine heterogeneity in predicted patent value by area of invention in Appendix Table E.2.

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. As noted above, the withdrawn applications were identified using the USPTO patent document pre-grant authority files.

- Six publication numbers are listed in the USPTO patent document pre-grant authority files but are not available in any of our other datasets;[30] we exclude these observations from our sample.

- Four publication numbers are missing from the Thomson Innovation database.[31] 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,[32] 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 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:

---

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

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

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

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

- Seven observations appear to be missing from Google's XML files of the published patent applications.[33] 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" field under Correspondence Address.

- The applications data contain 67 applications that were approved SIR (statutory invention registration) status but have the kind code "A1," instead of "H1" (as we would expect). We changed the kind code to "H1" for these applications, and they are therefore dropped from our sample.

A.3 Description of US Treasury tax files

All firm-level variables are constructed from annual business tax returns over the years 1997-2014: C-Corporations (Form 1120), S-Corporations (Form 1120S), and Partnerships (Form 1065). Worker-level variables are constructed from annual tax returns over the years 1999–2014: Employees (form W2) and contractors (form 1099).[34]

Variable Definitions

- Revenue

  - Line 1c of Form 1120 for C-Corporations, Form 1120S for S-Corporations, and Form 1065 for partnerships. When 1c is not available, we use 1a, which is gross receipts. We replace negative revenue entries, which are very rare, with missing values.

- 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.

- Total Deductions

  - For C-corporations, line 27 on Form 1120. For S-corporations, line 20 on Form 1120S. For partnerships, line 21 on Form 1065.

- Labor Compensation

  - For C-Corporations, sum of lines 12, 13, 24, and 25 on Form 1120.[35] For S-Corporations, sum of lines 7, 8, 17, and 18 for Form 1120S. For partnerships, sum of lines 9, 10, 18, and 19 on Form 1065. These lines are compensation to officers, salaries and wages, retirement plans, and employment benefit programs, respectively.[36]

- Value Added

---

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

[34]W2 data are not available in 1997–1999.

[35]Ideally, we could also add Schedule A line 3, which is the cost of labor on the COGS Form 1125-A, but these data are not available. However, the W2-based measure of compensation avoids this issue.

[36]For partnerships, the compensation to officers term is called "Guaranteed payments to partners."

56

- 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 3 on Form 1065.[37]

- 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, we define operating profits as the sum of lines 1c, 18, and 20, 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.

- EBITD

  - EBITD is total income less total deductions (other than interest and depreciation).
  - For C-Corporations, it is the sum of lines 11, 18, and 20, less 27 on Form 1120.
  - For S-Corporations, it is the sum of lines 1c, 13, and 14 less 20 on Form 1120S.
  - For partnerships, it is the sum of lines 1c, 15, and 16c less 21 on Form 1065.

- Employment

  - Number of W2s associated with an Employer Identification Number (EIN).

- Wage bill per worker

  - Sum of W2 box 1 payments divided by number of W2s for a given EIN.

- Surplus

  - Sum of EBITD and Wage bill, which is the sum of W2 box 1 payments for a given EIN.

- Inventor earnings per inventor

  - Wage bill per worker for workers who are identified as inventors by Bell et al. (2016).

- Cohort earnings per worker

  - Wage bill per worker for workers who were employed at the firm in the year of application regardless of whether or not they stay at the firm.

- Stayer earnings per worker

  - Stayers are cohort earning per worker for the set of workers who are still at the firm.

---

[37]Line 3 is calculated as line 1c minus line 2.

- Leaver earnings per worker

    - Leavers are cohort earning per worker for the set of workers in the initial cohort who are no longer at the firm, i.e., are no longer receiving a W2 associated with the original firm that applied for a patent.

- Earnings Gap Q4-Q1

    - Average earnings within quartile four and quartile one of a firm's wage distribution.

- Separators

    - The number of workers who left the EIN in the previous year.

- Entrants

    - The number of workers who joined the EIN relative to the previous year.

- 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, and Box A of Form 1065 for partnerships.

- 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. Below we describe 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, IN-CORPORATED," 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.[38] We then standardize common terms that take

---

[38]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.

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.[39] 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 $SNAME_{USPTO}$.

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, or 1065 filings over the years 1997–2014. 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 Section 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. However, there are some cases in which multiple EINs are associated with a given $SNAME_{Treasury}$. In those cases, we chose the EIN with the largest total income in the year of application in order to select the most economically active entity associated with that standardized name.[40]

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 b, 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 is unmatched. Because of computational limitations, we limit comparisons to cases in which both a and b start with the same letter. Therefore we will miss any matches that do not share the same first letter. 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

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

---

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

[40]For example, if two EINs shared the same standardized name $SNAME_{Treasury}$ but one EIN made 50 million in total income and the other showed three million in total income, we chose the EIN that earns 50 million.

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, Ravikumar, and Fienburg 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 $s = s_1 \ldots s_K$ and $t = t_1 \ldots t_L$ 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 metric 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, Ravikumar, and Fienburg (2003).

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

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

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 (2016), will likely become clearer after reviewing the following examples and discussing how we selected the tuning parameters (such as the closeness threshold θ).

Example

USPTO Assignee Name	Compustat Firm Name (best match)	Match Score
angiotech pharmaceuticals corp	angiotech pharmaceuticals	.9982
assg brooks justin	brooks resources corp	.5857
hewlett packard development corp	hewlett packard corp	.8482
huawei device corp	huatue electronics corp	.0013
matsushita electric works corp	matson corp	.0012
olympus corp	olympus capital corp	.9109
safety crafted solutions corp	safety first corp	.3862
sc johnson home storage corp	sc holdings corp	.5144

This table provides a small sample of candidate matches from a USPTO to Compustat match.[41]

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 would only accept 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 to workers, of which 286 (13%) had matches in the Compustat data.[42]

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 .91. 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

---

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

[42]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.

prefix doesn't improve performance, which makes sense given that we block by the first letter already.[43] Finally, the threshold match score of .91 shows that we should only consider names a match if they are very close by our similarity metric. With these parameters, Type I and II errors are each below six percent.

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 assignee-based merge to match to the same EIN as the Bell et al. (2016) inventor-based merge. However, the Bell et al. (2016) 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. (2016) 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.

---

[43]It is computationally infeasible to compare every single entity against every other, so we utilized first-letter blocking 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.

B Appendix: Skill upgrading

Using equation (3), we can write the labor supply of type q workers to firm j as:

$$\begin{aligned} N_j(q) &= N(q)(w_j(q) - b_j(q))^\beta \\ &= N(q)(\theta[T_j\mu_j q - b(q)])^\beta \end{aligned}$$

If $b(q) = cq^{1+r}$, then:

$$N_j(q) = N(q)(\theta[T_j\mu_j - cq^r]q)^\beta$$

For $q' > q$, we have

$$\ln\frac{N_j(q')}{N_j(q)} = \ln\frac{N(q')}{N(q)} + \beta\ln\frac{q'}{q} + \beta\ln\left(\frac{T_j\mu_j - cq'^r}{T_j\mu_j - cq^r}\right)$$

Note that when $r = 0$, this last term disappears, implying the skill mix is invariant to the state of firm productivity. However, when $r > 0$, we have $d\ln\frac{N_j(q')}{N_j(q)}/dT_j > 0$, implying firms upgrade their workforce in response to a productivity shock.

C Appendix: Poisson model of patent value

Recall that the probability mass function for a Poisson distributed outcome Y with mean λ can be written:

$$p(Y|\lambda) = \exp(Y\lambda - \exp(\lambda))/Y!$$

Let $Y_a = (Y_1, ..., Y_{m_a})$ and $X_a = (X_1, ..., X_{m_a})$ denote the vectors of outcomes and covariates respectively in an art unit a. Supposing $Y_j|X_a \stackrel{iid}{\sim} Poisson\left(X_j'\delta + v_a\right)$ where v_a is a scalar art unit effect, we can write:

$$
\begin{aligned}
\ln p(Y_a|X_a, v_a) &= \sum_{j=1}^{m_a} \ln p\left(Y_j|X_j'\delta + v_a\right) \\
&= \sum_{j=1}^{m_a} Y_j\left(X_j'\delta + v_a\right) - \exp\left(X_j'\delta + v_a\right) - \ln\left(Y_j!\right)
\end{aligned}
$$

The random effects Poisson likelihood of an art unit a can be written:

$$L(Y_a|X_a) = \frac{1}{\sqrt{2\pi}\sigma_\eta} \int \exp\left\{\ln p(Y_a|X_a) - \frac{1}{2}\frac{v^2}{\sigma_v^2}\right\} dv$$

By independence across art units, the full log likelihood can be written $\sum_a \ln L(Y_a|X_a)$.

The first order condition for the coefficient vector δ is:

$$
\begin{aligned}
\frac{d}{d\delta} \sum_a \ln L(Y_a|X_a) &= \sum_a \frac{\int \left(\sum_{j=1}^{m_a}\left[Y_j - \exp\left(X_j'\delta + v\right)\right]X_j\right)\exp\left\{\ln p(Y_a|X_a, v) - \frac{1}{2}\frac{v^2}{\sigma_v^2}\right\}dv}{\int \exp\left\{\ln p(Y_a|X_a, v) - \frac{1}{2}\frac{v^2}{\sigma_v^2}\right\}dv} \\
&= \sum_a \sum_{j=1}^{m_a} \left[Y_j - \int \exp\left(X_j'\delta + v\right)\omega_a(v)dv\right]X_j = 0
\end{aligned}
$$

where the weighting function $\omega_a(z) = \dfrac{\exp\left\{\ln p(Y_a|X_a, z) - \frac{1}{2}\frac{z^2}{\sigma_v^2}\right\}}{\int \exp\left\{\ln p(Y_a|X_a, v) - \frac{1}{2}\frac{v^2}{\sigma_v^2}\right\}dv}$ is the posterior density of v given the observables in art unit a. Note that this is a shrunken version of the usual Poisson orthogonality condition that is robust to misspecification of features of the conditional distribution other than the mean (Wooldridge 2010). The weights, however, rely on the exponential nature of the Poisson density function which, if misspecified, will yield inconsistency in small art units. In large art units, however, the posterior will spike around the "fixed effect" estimate of v, which is again robust to misspecification of higher moments of the conditional distribution.

The first order condition for the variance σ_v is:

$$
\begin{aligned}
\sum_a \frac{d}{d\sigma_v} \ln L(Y_a|X_a) &= \sum_a \left\{ -\frac{1}{\sigma_\eta} + \frac{\int \frac{\eta^2}{\sigma_\eta^3}\exp\left\{\ln p(Y_a|X_a, v) - \frac{1}{2}\frac{v^2}{\sigma_v^2}\right\}dv}{\int \exp\left\{\ln p(Y_a|X_a, v) - \frac{1}{2}\frac{v^2}{\sigma_v^2}\right\}dv} \right\} \\
&= \frac{1}{\sigma_\eta^3} \sum_a \left[\int \omega_a(v)v^2 dv - \sigma_v^2\right] = 0.
\end{aligned}
$$

This simply says that the posterior variance of v in each art unit should average across art units to σ_v^2.

D Appendix: $\sinh^{-1}(.)$ as a weighting of quantile treatment effects

Suppose Y is a continuous firm outcome (e.g., wage bill per worker) with potential outcomes $\{Y(0), Y(1)\}$ corresponding to whether a patent application is initially denied or allowed. Let $\{F_0(.), F_1(.)\}$ denote the distribution functions of these potential outcomes, and $q_d(\tau) = F_d^{-1}(.)$ the corresponding quantile functions for $d \in \{0, 1\}$. Note that we can write $Y(d) = q_d(U)$, where U is a uniform random variable.

Consider a monotone increasing concave function $g(.) : \mathbb{R} \to \mathbb{R}$ that is everywhere differentiable, an example of which is $\sinh^{-1}(.)$. The average treatment effect on the transformed outcome can be written

$$
\begin{aligned}
E\left[g(Y(1)) - g(Y(0))\right] &= E\left[g(q_1(U)) - g(q_0(U))\right] \\
&= \int_0^1 \left[g(q_1(\tau)) - g(q_0(\tau))\right] d\tau \\
&= \int_0^1 g'(\bar{q}(\tau))(q_1(\tau) - q_0(\tau)) d\tau \\
&= K \int_0^1 \omega(\bar{q}(\tau)) \underbrace{(q_1(\tau) - q_0(\tau))}_{QTEs} d\tau
\end{aligned}
$$

where the third line follows from the mean value theorem for a function $\bar{q}(\tau) \in [q_0(\tau), q_1(\tau)]$ that is monotone in τ. Note that the treatment effect on the τ quantile is $q_1(\tau) - q_0(\tau)$. The constant $K = \int \omega(\bar{q}(\tau)) d\tau$ is a conversion factor. When the treatment effects are uniformly small across quantiles, $K \approx E[g'(Y)]$. The term $\omega(\bar{q}(\tau)) = g'(\bar{q}(\tau))/K$ is a weighting function that, by concavity, is monotonically decreasing in τ. Hence, the concave transformation $g(.)$ downweights impacts at top quantiles.

E   Appendix: Additional figures and tables

Figure E.1: Years Until Initial Decision
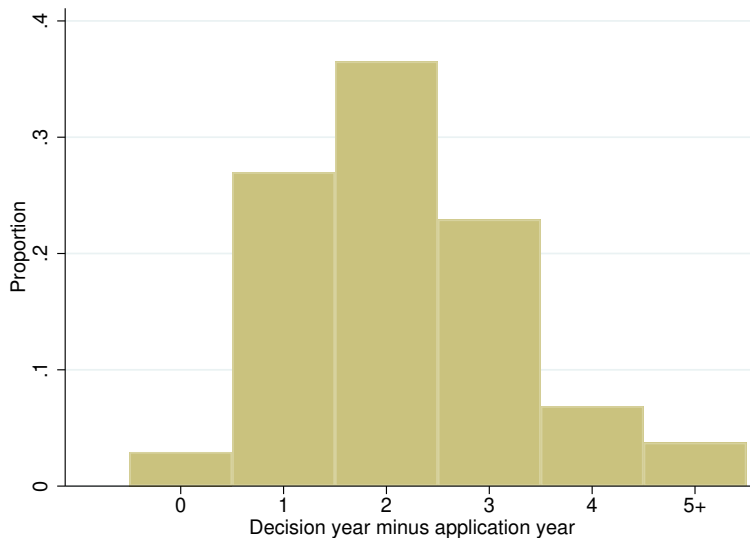
Notes: This figure plots a histogram of the years until the initial patent application decision for the sample of patent assignees by application pairs in the bottom row of Panel A of Table 1 (N=99,871).

Figure E.2: Years Until Patent Grant for Initially Rejected Patent Applications
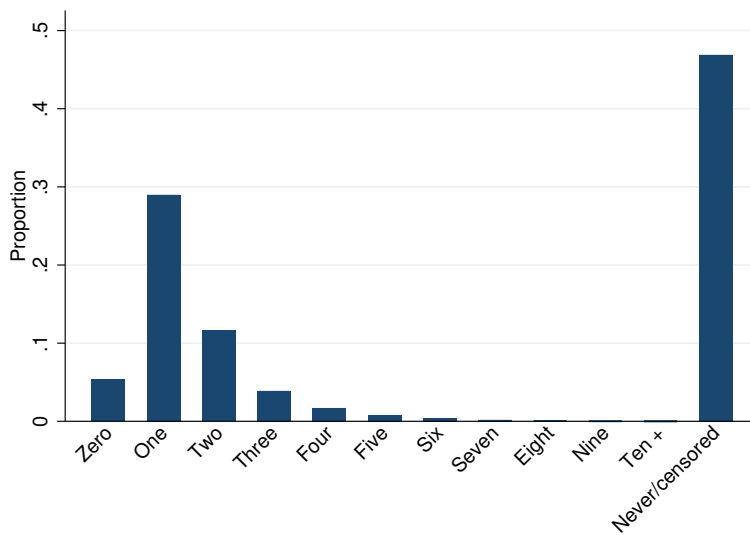
Notes: This figure plots a histogram of the years until a patent grant for the subsample of patent assignee by application pairs in the bottom row of Panel A of Table 1 (N=99,871) which receive an initial rejection (N=88,298).

Table E.1: Testing for Spatial Correlation in Initial Allowance Decisions

	Initially allowed				Top quintile			
	Analysis sample							
	State		Zip		State		Zip	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Intra-class correlation (ρ)	0.000	0.000	0.067	0.101	0.113	0.170	0.000	0.014
p-value	1.000	1.000	0.296	0.438	0.163	0.123	1.000	0.255
Observations	9,735	9,735	9,735	9,735	1,947	1,947	1,947	1,947
Geographic units	51	51	4,504	4,504	49	49	1,250	1,250
AU-AY FEs		✓		✓		✓		✓

Notes: This table reports the results of tests for whether initial patent allowances are geographically clustered, separately for the analysis sample and for the top-quintile predicted patent value sample. "Intra-class correlation (ρ)" reports the ratio of a random effects estimate of the geographic variance component to the sum of the geographic and idiosyncratic variance components. The p-value reports a Breusch-Pagan Lagrange multiplier test of the null hypothesis that $\rho=0$. AU-AY FEs denotes the inclusion of Art Unit (AU) by application year (AY) fixed effects.

Table E.2: Mean $\bar{\hat{\xi}}$ by Technology Center

Technology center	$\bar{\hat{\xi}}$	N	Technology center	$\bar{\hat{\xi}}$	N
Business Methods - Finance	22.356	96	Amusement & Education Devices	3.349	176
Electronic Commerce	14.034	245	Semiconductors, Circuits, & Optics	3.202	199
Databases & File Mgmt	11.924	203	Refrigeration & Combustion	3.074	265
Aero, Agriculture, & Weaponry	10.927	127	Telecomms: Analog Radio	3.052	36
2180: Computer Architecture	10.510	48	Mining, Roads, & Petroleum	3.050	502
Organic Compounds	10.272	41	Molec Bio & Bioinformatics	3.034	28
Tires, Adhesives, Glass, & Plastics	10.151	95	Static Structures & Furniture	2.957	264
Organic Chemistry	8.038	31	Microbiology	2.894	47
Combust & Fluid Power Systems	7.959	109	Business Methods	2.808	115
Selective Visual Display Systems	6.845	173	Computer Networks	2.580	51
Manufact Devices & Processes	6.709	234	Fuel Cells & Batteries	2.545	166
Optics	6.646	272	Software Development	2.262	68
Organic Chemistry & Polymers	6.312	96	2110: Computer Architecture	2.183	38
Computer Graphic Processing	6.139	286	Medical Instruments	2.018	73
Selective Communication	5.387	214	Semiconductors & Memory	1.876	160
Coating, Etching, & Cleaning	5.025	36	Metallurgy & Inorganic Chemistry	1.723	66
Memory Access & Control	4.876	48	Medical & Surgical Instruments	1.672	64
Surface Transportation	4.425	294	Radio, Robotics, & Nucl Systems	1.485	58
Electrical Circuits & Systems	4.198	255	Receptacles, Shoes, & Apparel	1.424	365
Multiplex & VoIP	4.052	22	Kinestherapy & Exercising	1.236	96
Misc. Computer Applications	3.879	100	Fluid Handling	0.704	188
Material & Article Handling	3.753	184	Chemical Apparatus	0.434	23
Graphical User Interface	3.487	103			

Notes: This table reports the mean predictions of ex-ante value $\hat{\xi}$ by USPTO technology center of the application; technology centers are administrative groupings of art units designated by the USPTO. The sample is observations from our analysis sample whose application belongs to a technology center with more than 20 observations in the analysis sample (N=6,402). $\hat{\xi}$ is measured in millions of 1982 USD.

Table E.3: Pass-Through Estimates: Three-Year Average of Surplus per Worker

	Wage bill per worker		Avg stayer earnings		Avg earnings of stayers minus earnings in appl yr		Avg non-inventor stayer earnings	
	(1a) OLS	(1b) IV	(2a) OLS	(2b) IV	(3a) OLS	(3b) IV	(4a) OLS	(4b) IV
Panel A: level								
Surplus per worker	0.20	0.34	0.26	0.66	0.25	0.60	0.21	0.60
	(0.01)	(0.15)	(0.01)	(0.34)	(0.01)	(0.31)	(0.01)	(0.28)
1st stage F		8.23		8.82		8.82		6.11
Exogeneity		0.298		0.182		0.212		0.070
Panel B: sinh$(\cdot)^{-1}$								
Surplus per worker	0.05	0.21	0.05	0.27	0.05	0.23	0.05	0.28
	(0.00)	(0.12)	(0.00)	(0.15)	(0.00)	(0.14)	(0.00)	(0.16)
1st stage F		5.12		5.80		5.80		4.76
Exogeneity		0.076		0.075		0.119		0.052
Observations	83,230	83,230	81,093	81,093	81,093	81,093	77,957	77,957

Notes: This table reports OLS and IV estimates of the effect of increases in surplus per worker on selected earnings outcomes using three-year averages of surplus per worker. The excluded instrument is the interaction of top quintile of ex-ante value $\hat{\xi}$ category with a post-decision indicator and an indicator for the application being initially allowed. Controls include main effect of value category interacted with post-decision indicator and interaction of lower quintile value category with a post-decision indicator times an indicator for initially allowed, firm fixed effects, and Art Unit by application year by calendar year fixed effects. Standard errors (reported in parentheses) are two-way clustered by (1) art unit, and (2) application year by decision year. Panel B takes sinh$^{-1}(\cdot)$ transformations of the outcome and endogenous variables. "Exogeneity" reports p-value for test of null hypothesis that IV and OLS estimators have same probability limit. Stayers are defined as those who were employed by the same firm in the year of application. Surplus is EBITD (earnings before interest, tax, and depreciation) + wage bill. Earnings, wage bill, and surplus are measured in thousands of 2014 USD.