DISCUSSION PAPER SERIES

IZA DP No. 14082

The Impact of Regulation on Innovation

Philippe Aghion
Antonin Bergeaud
John Van Reenen

JANUARY 2021

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14082

The Impact of Regulation on Innovation

Philippe Aghion
Collège de France and LSE

Antonin Bergeaud
Banque de France and CEP

John Van Reenen
LSE, MIT and IZA

JANUARY 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Impact of Regulation on Innovation*

Does regulation affect the pace and nature of innovation and if so, by how much? We build a tractable and quantifiable endogenous growth model with size-contingent regulations. We apply this to population administrative firm panel data from France, where many labor regulations apply to firms with 50 or more employees. Nonparametrically, we find that there is a sharp fall in the fraction of innovating firms just to the left of the regulatory threshold. Further, a dynamic analysis shows a sharp reduction in the firm's innovation response to exogenous demand shocks for firms just below the regulatory threshold. We then quantitatively fit the parameters of the model to the data, finding that innovation at the macro level is about 5.4% lower due to the regulation, a 2.2% consumption equivalent welfare loss. Four-fifths of this loss is due to lower innovation intensity per firm rather than just a misallocation towards smaller firms and lower entry. We generalize the theory to allow for changes in the direction of R&D, and find that regulation's negative effects only matter for incremental innovation (as measured by citations and text-based measures of novelty). A more regulated economy may have less innovation, but when firms do innovate they tend to "swing for the fence" with more radical (and labor saving) breakthroughs.

JEL Classification: O31, L11, L51, J8, L25

Keywords: innovation, regulation, patents, firm size

Corresponding author:
John Van Reenen
MIT
77 Massachusetts Ave
Cambridge, MA 02139
USA
E-mail: vanreene@mit.edu

* This project began with a collaboration with Costas Meghir who was instrumental in our thinking over the paper and gave tremendous inspiration and help with framing the theory and empirical approach. Matthieu Lequien generously allowed us to use his data matching and cleaning patents, which was a huge multi-year project. Heidi Williams inspired us to push the work substantially further from an earlier draft. Great comments have come from Daron Acemoglu, Ufuk Akcigit, Hunt Allcott, Sharon Belenzon, David Demming, Bernhard Ganglmair, Gordon Hansen, Claire Lelarge, Cyril Verluise, Gianmarco Ottaviano, Mike Webb and participants in seminars in Bath, Boston College, CEPR Globalization and Labor workshop, EFF Milan, Harvard, Microsoft Research Group, NBER, Peking HSBC, San Diego and UCL. Financial support from the ESRC, Sloan Founda- tion, SRF and Schmidt Sciences is gratefully acknowledged. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-EQPX-17 - Centre d'accès securisé aux données – CASD)

1 Introduction

There is considerable literature on the economic impacts of regulations, but relatively few studies on their impact on technological innovation. Most analyses focus on the static costs (and benefits) of regulation rather than on its dynamic effects. Yet these potential growth effects are likely to be much more important in the long-run. Harberger triangles may be small, but rectangles can be very large. Many scholars have been concerned that slower growth in countries with heavy labor regulation, could be due to firms being reluctant to innovate due to the burden of red tape. For example, the slower growth of Southern European countries and parts of Latin America have often be blamed on onerous labor laws.[1]

Identifying the innovation effects of labor regulation is challenging. The OECD, World Bank, IMF and other agencies have developed various indices of the importance of these regulations, based on examination of laws and surveys of managers. These indices are then often included in econometric models and sometimes found to be significant. Unfortunately, these macro indices of labor law are correlated with many other unobservable factors that are hard to convincingly control for.[2] To address this issue, we exploit the fact that many regulations are size contingent and only apply when a firm gets sufficiently large. In particular, the burden of French labor legislation substantially increases when firms employ 50 or more workers. For example, such firms must create a works council with a minimum budget of 0.3% of total payroll, establish a health and safety committee, appoint a union representative and so on (see Appendix A for more institutional details). Several authors have found that these regulations have an important effect on the size of firms. Indeed, unlike the US firm size distribution, for example, in France, there is a clear bulge in the number of firms that are just below this regulatory threshold.[3]

Existing models that seek to rationalize these patterns have not usually considered how this regulation could affect innovation, as technology has been assumed exogenous. But when firms are choosing whether or not to invest in innovation, regulations are also likely to matter. Intuitively, firms may invest less in R&D as there is a very high cost of growing if the firm crosses the regulatory threshold. In the first part of the paper we formalize this intuition

[1] See for example, Gust and Marquez, 2004; Bentolila and Bertola, 1990, Bassanini et al., 2009, Schivardi and Schmitz, 2020).

[2] Furthermore, it may be that the more innovative countries are less likely to adopt such regulations (e.g. Saint-Paul, 2002).

[3] See Garicano et al., 2016; Gourio and Roys, 2014; Ceci-Renaud and Chevalier, 2011; and Smagghue, 2020. Often, it is hard to see such discontinuities in the size distribution at regulation thresholds (e.g. Hsieh and Olken, 2014). A reason for the greater visibility in France is because the laws are more strictly enforced through large numbers of bureaucratic enforcers and strong trade unions.

using a simple version of the Klette and Kortum (2004) model of growth and firm dynamics, with discrete time and two-period lived individuals (but potentially infinitely lived firms). Our model delivers a number of predictions regarding the shape of the equilibrium relationship between innovation and firm size and the overall firm size distribution. In particular we obtain the intuitive prediction that the regulatory threshold discourages innovation most strongly for firms just below the threshold, although it also discourages and shallows the innovation-size gradient for all firms larger than the threshold. This is because the growth benefits of innovation are lower due to the implicit regulatory tax.

We use the discontinuous increase in cost at the regulatory threshold to test the theory in two ways when taking it to our rich panel data on the population of French firms. First, we investigate non-parametrically how innovation changes with firm size. As expected there is a sharp fall in the fraction of innovative firms just to the left of the regulatory threshold, an "innovation valley" that is suggestive of a chilling effect of the regulation on the desire to grow. Moreover, there is a flattening of the innovation-size relationship to the right of the threshold, consistent with a greater tax on growth. Although the cross sectional evidence is suggestive, there could be many other reasons why firms are heterogeneous near the regulatory threshold, so we turn to a second and stronger test by exploiting the panel dimension of our data. Specifically, based on wide class of models that predict that an increase in market size should have a positive effect on innovation (e.g. Acemoglu and Linn, 2004), we analyze the heterogeneous response of firms with different sizes to exogenous demand shocks. We use a shock based measure based on changes in growth in export product markets (disaggregated HS6 products by country destination) interacted with a firm's initial distribution of exports across these export markets (see Hummels et al., 2014 or Mayer et al., 2016) We first show that these positive market size shocks significantly raise innovative activity. We then examine the heterogeneity in firm responsiveness to these demand shocks depending on (lagged) firm size. We find a sharp reduction in firms' innovation response to the shock for firms with size just below the regulatory threshold. Consistent with intuition and our simple model, firms appear reluctant to take advantage of exogenous market growth through innovating when they will be subject to a wave of labor regulation.

Having established that the qualitative implications of the model are consistent with the data, we use the structure of our model (and empirical moments of the data) to quantitatively estimate the impact of the regulation on aggregate innovation and welfare. Our baseline estimates suggest that the regulation is equivalent to a tax on profit of about 2.5% that reduces aggregate innovation by around 5.4% (equivalent to cutting the growth rate from say, 2 to 1.9

percentage point per annum) and reduces welfare by at least 2.2% in consumption equivalent terms. This is partly through misallocation from lowering entry and shifting the size distribution downwards, but the vast majority of this aggregate impact is through lower innovation per firm once they reach a certain size. This implies that the existing structural static analyses of the output loss have significantly underestimated the cost of the regulation.

One caveat to our welfare conclusions comes from generalizing the model to allow firms to invest in a mixture of radical and incremental innovation. We find that the regulation deters incremental R&D, but if a firm is going to innovate it will try to "swing for the fence" to avoid being only slightly to the right of the threshold. Measuring radical innovation by either future citations or a machine learning approach based on novelty in the patent text, we find that the negative effects of regulations are confined to incremental patents. Similarly, we find the regulation biases innovation towards automative labor-saving patents.

Related literature

Our paper relates to several strands of literature. More closely related to our analysis are papers that look at the effects of labor laws regulations on innovation. In Acharya et al. (2013a) higher firing costs reduce the risk that firms would use the threat of dismissal to hold their employees' innovative investments up. They find evidence in favor of this using macro time series variation on Employment Protection Law (EPL) for four OECD countries. Acharya et al. (2013b) also finds positive effects using staggered roll out of employment protection across US states.[4] Griffith and Macartney (2014) use multinational firms patenting activity across subsidiaries located in different countries with various levels of EPL.[5] Using this cross sectional identification, they find that radical innovation was negatively effected by EPL, but incremental innovations were not.[6] By contrast, Alesina et al. (2018) find that less regulated countries have larger high tech sectors. All of these papers use macro (or at best, state-level) variation whereas we focus on cross firm variation. Garcia-Vega et al. (2019) analyze a reform that relaxed a size contingent labor regulation in Spain and find an increase in innovation. Our empirical results are consistent with this, but we go beyond the analysis in this paper by developing a model of

[4]This is the same empirical variation used by Autor et al. (2007) who actually found falls in TFP and employment from EPL. And Bena et al. (2020) finds a positive impact on process innovation using the same design.

[5]See also Cette et al. (2016) who document a negative effect of EPL on capital intensity, R&D expenditures and hiring of high skill workers. More generally, Porter and Van der Linde (1995) argue that some regulations, such as those to protect the environment, can have positive effects on innovation.

[6]Note that this is the opposite of what we find using our within country identification. Labor regulation discourages low value innovation, but has no impact on high value innovation.

labor regulation and innovation with endogenous firm size distribution, that is matched with the data to obtain structural parameters, enabling us to perform aggregate counterfactuals.

Second, several structural papers look at the effects of labor regulations on employment and welfare, in particular Braguinsky et al. (2011) on Portugal, Gourio and Roys (2014) and Garicano et al. (2016) on France. However, these papers do not allow for endogenous innovation. More generally, there is a large literature focusing on how various kinds of distortions can affect aggregate productivity through the resulting misallocation of resources away from more productive firms and towards less productive firms. As Restuccia and Rogerson (2008) and Parente and Prescott (2000) have argued, these distortions imply that more efficient firms produce too little and employ too few workers. Hsieh and Klenow (2009) show that the resulting misallocation accounts for a significant fraction of the differences in aggregate productivity between the US, China and India and Bartelsman et al. (2013) confirm this finding using micro data from OECD countries.[7] Boedo and Mukoyama (2012) and Da-Rocha et al. (2019) have shown firing costs hinder job reallocation and reduce allocative efficiency and aggregate productivity. The additional effect of barriers to reallocation when productivity is endogenous is also the focus of Gabler and Poschke (2013), Da-Rocha et al. (2019), and Bento and Restuccia (2017). Samaniego (2006) highlights the effects of firing costs in a model with productivity growth. He considers, however, only exogenous productivity growth and studies how the effects of firing costs differ across industries. Poschke (2009) is one of the few exceptions that studies the effects of firing costs on aggregate productivity growth. Mukoyama and Osotimehin (2019) is perhaps the most closely related paper to ours and finds a negative growth effect of the firing tax equivalent to a 5% labor tax (in the entrant-innovation model in the US) in a calibrated aggregate model with endogenous innovation. Unlike our approach, their paper does not have closed form solutions for the policy rules with taxes so have to rely on simulation methods. We contribute to this part of the literature by introducing an explicit source of distortion, namely the regulatory firm size threshold that goes beyond just firing costs, and by looking at how this regulation interacts with exogenous market size shocks using firm-level micro-econometric analysis.

Third, a body of work looks at the effects of EPL on the adoption of new technologies (e.g. Manera and Uccioli, 2020), especially information and communication technology. For

[7]In development economics many scholars have pointed to the "missing middle" , i.e. a preponderance of very small firms in poorer countries compared to richer countries (see Banerjee and Duflo, 2005, or Jones, 2011). Besley and Burgess (2000) suggest that heavy labor regulation in India is a reason why the formal manufacturing sector is much smaller in some Indian states compared to others.

example, Bartelsman et al. (2016) argue that risky technologies require frequent adjustments of the workforce. By increasing the costs of such adjustments, EPL will deter technology adoption. Similarly Samaniego (2006) finds that EPL slows diffusion and Saint-Paul (2002) finds a smaller share of the economy in risky sectors when EPL are strong. Our approach is different as it focuses on technological innovation at the frontier rather than the adoption of existing technologies. Unlike emerging economies, advanced countries such as the US or France cannot rely solely on catch-up diffusion for long-run sustainable growth.

Fourth, our paper is related to public finance as we model regulation as an implicit tax, and a number of papers have examined how personal and business taxes affect innovation (see Akcigit and Stantcheva, 2020, for a recent survey). Like us, other tax papers use nonlinearities to identify behavioral parameters (e.g. Saez, 2010; Chetty et al., 2011; Kleven and Waseem, 2013; Kaplow, 2013 and Aghion et al., 2019b) and we contribute to this literature by bringing labor regulations, innovation and patenting into the picture.[8]

Fifth, there is an older literature looking at one form on labor regulation - union power - on innovation.[9] These papers found ambiguous theoretical and empirical effects . Finally, the heterogeneous effects of demand shocks on types of innovation is also a theme in the literature of the effects of the business cycle on innovation (Schumpeter, 1939; Shleifer, 1986; Barlevy, 2007). Recent work by Manso et al. (2019) suggests that large positive demand shocks (booms) generate more R&D, but this tends to ”exploitative” (incremental) rather than "exploratory" (radical) innovation. We find that the impact of regulation following a demand shock discourages incremental (but not radical) innovation.

The structure of the paper is as follows. Section 2 develops a simple model of how innovation can be affected by size-contingent regulation. Section 3 confronts the main qualitative predictions of the model to the data, using both a non-parametric cross sectional analysis and a dynamic analysis of the response to shocks. Section 4 uses the theory and empirical moments to estimate the equilibrium effect of regulation on aggregate innovation and welfare. Section 5 presents a number of theoretical and empirical extensions and robustness tests, most importantly allowing for radical and incremental innovation. Section 6 concludes. In Online

[8]This is important as Hopenhayn (2014) has argued that tax-driven reallocation distortions typically have only second order welfare effects unless there is rank reversal. Changing innovation is potentially a way of generating larger negative welfare effects that goes beyond static models.

[9]See Menezes-Filho et al. (1998) for a survey and evidence. The common view is that the risk of ex post hold-up by unions reduces innovation incentives (Grout, 1984). But if employees need to make sunk investments there could be hold up by firms (this is the intuition of the Acharya et al., 2013a,b papers).

Appendices, we present institutional details of the labor regulations (A), data details (B), further theoretical results (C) and additional econometric exercises (D).

2 Theory

In this section we present our basic theory built around a simplified Aghion et al. (2018b) set up where we introduce size contingent regulations. This enables us to analytically characterize firms' innovation decisions depending on their size and the regulation. We close the model by solving for the steady state firm size distribution incorporating both incumbent growth and entry/exit dynamics. We show how firm and economy wide innovation and size change with the stringency of the regulation. Throughout, we explore what the model implies for the steady state joint distributions of innovation and employment as well as how firms should respond to the exogenous demand shocks we will exploit in the empirical section.

2.1 A simplified Klette-Kortum model

We consider a simple discrete time version of the Schumpeterian growth model with firm dynamics by Klette and Kortum (2004) where firm owners live for only two periods. This two-period specification is drawn from Aghion et al. (2018b) and simplifies exposition of the model. We consider the case of infinitely lived owners in subsection 5.3, which delivers similar results. In the first period of her life, a firm owner decides how much to invest in R&D. In the second period, she chooses labor inputs, produces and realizes profits. At the end of the period, her offspring inherits the firm at its current size and a new cycle begins again.[10]

We assume that individuals have intertemporal log preferences:

$$U = \sum_{t>0} \beta^t \log(C_t), \tag{1}$$

associated with a budget constraint:

$$w_t + (1 + r_t)a_t = a_{t+1} + C_t,$$

where w_t is the wage at time t, C_t is consumption, and a_t is an asset that yields an interest rate r_t, . This immediately gives the Euler equation: $\beta(1 + r_t) = 1 + g_t$. We consider the economy

[10]We do not consider bequest motives, but the extension to infinitely lived agents implicitly encompasses this incentive.

on a balanced growth path where final output y and consumption grow at a constant rate which we denote by g, so that the Euler equation can be expressed as $\beta = \frac{1+g}{1+r}$, where β is the discount factor and r is the steady-state level of interest rate. There is a continuous measure L of production workers, and a mass 1 of intermediate firm owners every period. Each period the final good is produced competitively using a combination of intermediate goods according to the production function:

$$\ln y = \int_0^1 \ln(y_j)dj,$$

where y_j is the quantity produced of intermediate j. Intermediates are produced monopolistically by the firm who innovated last within that product line j, according to the linear technology $y_j = A_j l_j$ where A_j is the product-line-specific labor productivity and l_j is the labor employed for production. This implies that the marginal cost of production in j is simply w/A_j. A firm is defined as a collection of production units (or product lines/varieties) and expands in product space through successful innovation.

To innovate, an intermediate firm i combines its existing knowledge stock that it accumulated over time (n_i, the number of varieties it operates in) with its amount of R&D spending (R_i) according to the following Cobb-Douglas knowledge production function:

$$Z_i = \left(\frac{R_i}{\zeta y}\right)^{\frac{1}{\eta}} n_i^{1-\frac{1}{\eta}}, \tag{2}$$

where Z_i is the Poisson innovation flow rate, η is a concavity parameter and ζ is a scale parameter. This generates the R&D cost of innovation: $C(z_i, n_i) = \zeta n_i z_i^\eta y$, where $z_i \equiv Z_i/n_i$ is simply defined as the innovation intensity of the firm.

When a firm is successful in its current R&D investment, it innovates over a randomly drawn product line $j' \in [0, 1]$. Then, the productivity in line j' increases from A'_j to $A'_j\gamma$ and the firm becomes the new monopoly producer in line j' and thereby increases the number of its production lines to $n_i + 1$. At the same time, each of its n_i current production lines is subject to the risk of being replaced by new entrants and other incumbents (a creative destruction probability that we denote x). Thus the number of production units of a firm of size n_i increases to $n_i + 1$ with probability $Z_i = n_i z_i$ and decreases to $n - 1$ with probability $n_i x$. A firm that loses all of its product lines exits the economy.

Because of the Cobb-Douglas aggregator, the final good producer spends the same amount y on each variety j. As a result, final good production function generates a unit elastic demand with respect to each variety: $y_j = y/p_j$. Combined with the fact that firms in a single product line compete *a la* Bertrand, this implies that a monopolist with marginal cost w/A_j will follow limit pricing by setting its price equal to the marginal cost of the previous innovator $p_j = \gamma w/A_j$.[11]

The resulting equilibrium quantity and profit in product line j are:

$$y_j = \frac{A_j y}{\gamma w} \text{ and } \Pi_j = \left(1 - \frac{1}{\gamma}\right) y,$$

and the demand for production workers in each line is given by $y/(\gamma w)$. Firm i's employment is then equal to its total manufacturing labor, aggregating over all n_i lines where i is active, N_i. Namely:

$$L_i = \int_{j \in N_i} \frac{y}{w\gamma} dj = \frac{y n_i}{w\gamma} = \frac{n_i}{\omega\gamma}, \tag{3}$$

where $\omega = w/y$ is the output-adjusted wage rate, which is invariant on a steady state growth path. Importantly for us, a firm's employment is strictly proportional to its number of lines n_i.

2.2 Regulatory threshold and innovation

We model the regulation by assuming that a tax on profit must be incurred by firms with a labor force that is greater than a given threshold $\bar{l}$ (50 in our application in France). We suppose that $\bar{l}$ is sufficiently large that entrants do not incur this tax upon entry. There corresponds a cutoff number of varieties $\bar{n} = \bar{l}\omega\gamma$ to the employment threshold $\bar{l}$, such that if $n_i > \bar{n}$ profit is taxed at some additional positive marginal rate τ whereas the firm avoids this additional tax if $n_i \leq \bar{n}$.[12] Because firm owners live only for two periods, they can only expand the number of varieties of the firm by one extra unit during their lifetime. Hence, all the firms that start out with size $n_i < \bar{n} - 1$ or $n_i \geq \bar{n}$ act exactly as if the regulatory threshold did not exist. For firms that start with $n = \bar{n} - 1$, there is an additional cost to expanding by one extra variety.

[11] We implicitly assume a competitive fringe of firms with access to the previous technology in each sector; and that entering the market involves an ε entry cost. Then, as long as the new innovator sets a price which is less than the limit price equal to the marginal cost of fringe firms, no fringe firm will pay the entry cost. On the other hand, if the new innovator sets a price which is higher than the limit price, then she risks losing the market to a fringe firm.

[12] Unlike in Aghion et al. (2014) where the innovation cost is modelled as a labor cost, here innovation uses the final good y as an input. With labor as R&D input, total employment is $L_i = \frac{n_i}{\omega\gamma} + \zeta n_i z_i^\eta$, and thus varies with innovation rather than being proportional to n_i. We consider this extension in subsection 5.4 where R&D is labor. Increased R&D will then affect the equilibrium wage.

The owner of a n-size firm therefore maximizes their expected net present value over its innovation intensity $z \geq 0$:[13]

$$\max_{z \geq 0} \left\{ n\pi(n)y - \zeta n z^{\eta} y + \frac{1}{1+r} \mathbb{E}\left[n'\pi(n')y' \right] \right\},$$

where y' and n' denotes period 2's values for y and n_i and r is the interest rate. Dividing by y/n and using the fact that $\beta = (1+g)/(1+r)$, the above maximization problem can be re-expressed as:

$$\max_{z_i \geq 0} \left\{ \pi(n) - \zeta z^{\eta} + \beta z[(n+1)\pi(n+1) - n\pi(n)] + \beta x[(n-1)\pi(n-1) - n\pi(n)] \right\},$$

where $\pi(n) = \left(1 - \frac{1}{\gamma}\right)$ if $n < \bar{n}$ and $\pi(n) = \left(1 - \frac{1}{\gamma}\right)(1-\tau)$ if $n \geq \bar{n}$.

The intuition behind this equation is straightforward. The first term, $\pi(n)$ represents the gross flow profits per line today and the second term is the cost of research, ζz^{η}. The third term, $\beta z[(n+1)\pi(n+1) - n\pi(n)]$, is the (discounted) incremental profit gain tomorrow multiplied by the probability the firm innovates and thereby operates one more product line. The final term, $\beta x[(n-1)\pi(n-1) - n\pi(n)]$ is the (discounted) incremental profits loss per line tomorrow if the firm gets replaced in one of its product lines by a rival firm.

Whenever positive, the optimal innovation intensity is therefore given by:

$$z(n) = \begin{cases} \left(\dfrac{\beta(\gamma-1)}{\gamma\zeta\eta} \right)^{\frac{1}{\eta-1}} & \text{if } n < \bar{n} - 1 \\ \left(\dfrac{\beta(\gamma-1)(1-\tau\bar{n})}{\gamma\zeta\eta} \right)^{\frac{1}{\eta-1}} & \text{if } n = \bar{n} - 1 \\ \left(\dfrac{\beta(\gamma-1)(1-\tau)}{\gamma\zeta\eta} \right)^{\frac{1}{\eta-1}} & \text{if } n \geq \bar{n} \end{cases} \tag{4}$$

Much of the core economics of the paper can be seen in equation (4). Innovation intensity, $z(n)$, is highest for small firms a long way below the threshold (first row on right hand side of (4)), second highest for large firms over the threshold (third row) and lowest for middle sized firms just to the left of the threshold (middle row).

What we observe in the data is total innovation (as proxied by patent counts) which is $Z(n) = nz(n)$. Since employment is directly proportional to the number of product lines,

[13] Since we have shown that innovation per line is the same for firms of given size, we drop the firm i subscripts from here onwards for notational simplicity unless needed.

this implies that the slope of the innovation-size relationship will depend upon whether the firm lies above or below the regulatory threshold. Typically, the upwards sloping relationship between innovation and firm size should be steeper for small firms than for large firms and should fall and flatten discontinuously at the threshold. Furthermore, the ratio of the slopes of the innovation-size relationship for large versus small firms, relates directly to the underlying parameters of the model, and in particular upon the regulatory tax.[14] We will use this fact to empirically identify the magnitude of the regulatory tax, which we then use in our estimates of the aggregate impact of the regulation on innovation.

2.3 Regulatory threshold and firm size distribution

We now characterize the steady state distribution of firm size and look at how this distribution is affected by the regulatory tax. Let $\mu(n)$ be the share of firms with n lines. We first have a steady state condition saying that the number of exiting firms equals the number of entering firms in steady-state, namely: $\mu(1)x = z_e$, where z_e is the innovation intensity of entrants, which is the same as the probability of entry. Since x is the rate of creative destruction for any line, the number of exiting firms is therefore given by $\mu(1)x$.

For all $n > 1$, the steady state condition is that outflows from being a size n firm is equal to the inflows into becoming a size n firm. This can be expressed as:

$$n\mu(n)\left(z(n)+x\right) = \mu(n-1)z(n-1)(n-1) + \mu(n+1)x(n+1) \tag{5}$$

We know $z(n)$ for each n from equation (4) but we need to find the two remaining endogenous objects z_e and x. We close the model by considering the following two equations. First, the definition of μ gives $\sum_{n=1}^{\infty}\mu(n) = 1$. Second, the rate of creative destruction on each line is equal to the rate of creative destruction by an entrant plus the weighted sum of the flow probabilities $z(n)$ of being displaced by an incumbent of size n, namely:

$$x = z_e + \sum_{n=1}^{\infty}\mu(n)nz(n)$$

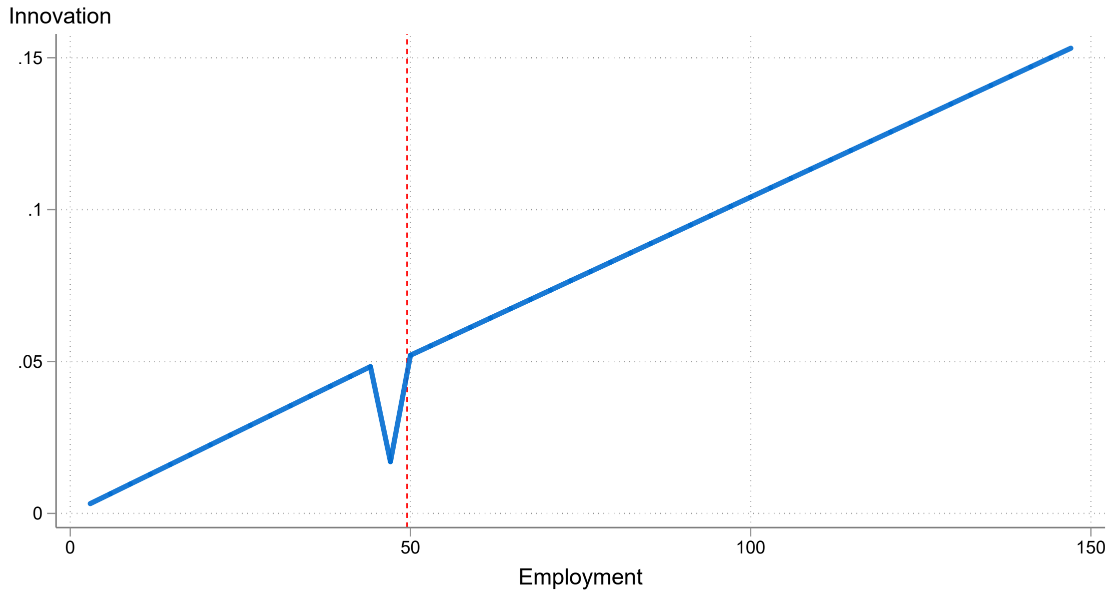
2.4 Solving the model

In Appendix C we detail how we solve the model numerically. The unknowns are $\mu(n)$ and $z(n)$ for all values of n as well as x and z_e, and the equations are those derived above. To

[14] The ratio of the innovation intensity of the first to third row in (4) is $(1-\tau)^{1/(1-\eta)}$. This can be empirically recovered from the relative slopes of the patents to size relationship before and after the regulatory threshold.

illustrate the effects, we first show firm-level innovation $Z(n) = z(n)n$ as a function of the firm's employment size $L = n/(\omega\gamma)$ in Figure 1. We see that firm-level innovation increases linearly with firm size until the firm nears the regulatory threshold, at which point there is a sharp innovation valley. After this, innovation again increases with firm size once the firm passes the threshold.

Figure 1: Total Innovation by firms of different employment sizes



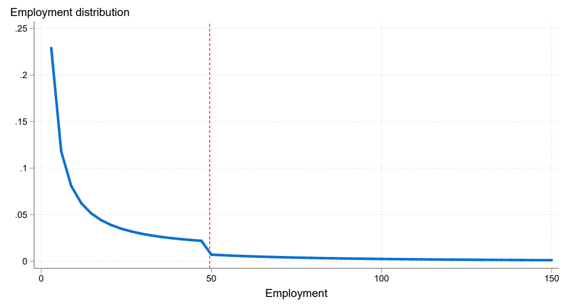
Notes: This is the total amount of innovation ($Z(n)$) by firms of different sizes (employment, $L = n/(\gamma\omega)$) by aggregating innovation intensities $z(n)$ across all its product lines (n) according to our baseline theoretical model. The y-axis is the corresponding value of total innovation Z. We use our baseline calibration values of $\tau = 0.025$, $\gamma = 1.3$, $\eta = 1.5$, β $\zeta = 1.65$ and $\omega = 0.28$ for illustrative purposes (see section 4 for a discussion).

In Figure 2 we plot the equilibrium firm size distribution, i.e. the value of the density $\mu(n)$ for each level of firm employment. Panel (a) uses a linear scale, but because the distribution nonlinear we plot it on a log-log scale in Panel (b) where it is broadly log-linear (the well-know power law as documented by Axtell, 2001 and many others). Note the departure from the power law around the regulatory threshold. The distribution bulges a bit as firms approach 50 and then discontinuously drops before falling again once firms pass the threshold. Unlike the innovation-size discontinuity, this "broken power law" in the French size distribution has been noted before in the literature (e.g. Ceci-Renaud and Chevalier, 2011), but the shape has proven difficult to fully rationalize in a model without endogenous innovation.[15]

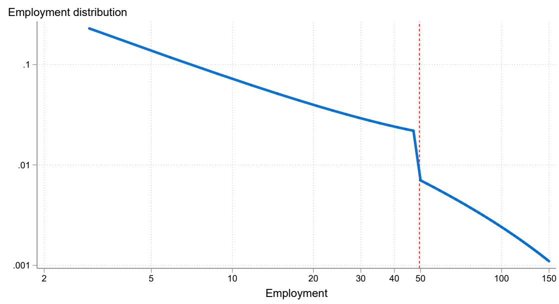
[15] In particular, although a purely static model like Lucas (1978) with regulation can rationalize a discontinuity at 50 and a downwards shift of the line, there should be no firms of size 50 and no bulge at 48 (firms just fully shift to avoid the regulation and spike at 49). Garicano et al., 2016 had to introduce ad hoc measurement error to rationalize the smoother bulge we see in the data around 45-50. This bulge (and the positive mass

Figure 2: Distribution of firm size ($\mu(n)$)

(a) Linear scale



(b) Log scale



Notes: These figures plot the density of firm employment, $\mu(n)$ according to our baseline theoretical model. Panel (a) uses a linear scale and Panel (b) uses a log-log scale. The calibration values are the same as Figure 1.

Although we took particular values of the parameters for illustrative purposes in Figures 1 and 2, these patterns are the same for any value of the regulatory tax (τ).[16] To see how τ qualitatively impacts the innovation-firm size relationship and the firm size distribution, we compare our results (solid blue) to an unregulated economy (i.e. $\tau = 0$, dashed red) in Figures 3(a) and 3(b). Four points are worth emphasizing. First, as expected, we observe no innovation valley when we remove the regulation Figures 3(a). Second, the level of innovation when $\tau > 0$ is lower than when $\tau = 0$ even for large firms to the right of threshold. This stems from the fact that the tax reduces innovation intensity even for these firms. Third, the total innovation gap between the regulated and unregulated economy gets larger as firm size increases because bigger firms have more product lines and the innovation intensity of each line is lower than that of small firms. This can be seen from (4), which showed that the slope of the line after the threshold is flatter than that for small firms with $n < \bar{n} - 1$. Fourth, in terms of the size distribution in Figure 3(b), we see that moving from $\tau = 0$ to $\tau > 0$ increases the share of firms that are below 50 employees and decreases the fraction of large firms. The regulation also generates a larger mass of firms just below the threshold as these firms choose not to grow in order to avoid getting hit by the regulatory tax.

We now put together all the effects of regulation together to compute the overall impact of regulation on the economy-wide innovation, $\mathcal{Z}(\tau) = \sum_{i=1}^{\infty} \mu(i)z(i)i + z_e$. Figure 4 shows the fall in total innovation in the regulated economy compared to the counterfactural unregulated economy (where we normalize aggregate innovation at 1). The magnitude of the fall in inno-

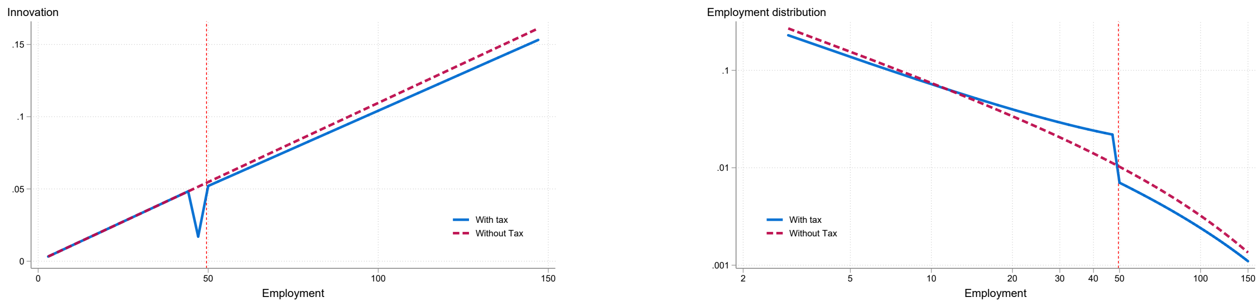
at 50) emerges more naturally with our dynamic endogenous innovation model.

[16] From equation (4), we know that we can take τ to lie anywhere between 0 and $1/\bar{n}$ in order to have an interior solution for $z(\bar{n})$.

Figure 3: Innovation and firm size distribution, with and without regulations

(a) Innovation

(b) Size distribution (log scale)



Notes: The blue solid line in this Figure reproduces Figure 1 in Panel (a) and Figure 2(b) in Panel (b). The red dashed line is for an unregulated economy with all the same parameters in the regulated economy except $\tau = 0$.

vation is clearly increasing in the intensity of the regulatory tax, τ. For example, there is a reduction in total innovation of 4% if $\tau = 0.02$ instead of zero. This fall in aggregate innovation comes from three sources. First, for a given firm size, the tax increase has a strong negative effect on innovation for firms just to the left of the threshold, and a smaller negative effect on innovation for all firms to the right of the threshold. Second, the tax increase reduces the mass of large firms, which are also the firms that do more innovation. Third, there is a fall in creative destruction as a higher tax means less entry (since all entrants need to innovate to displace an incumbent). When we use our data to quantify the model, we will decompose the fall of aggregate innovation into these different elements and show that the first element (incumbent innovation) dominates.

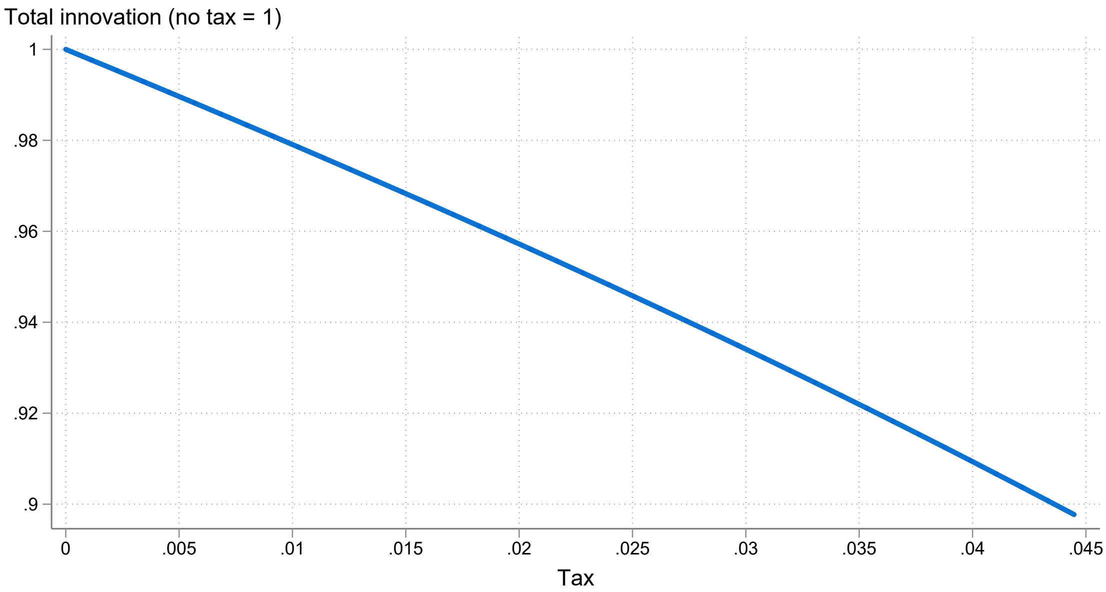
2.5 Effect of a demand shock

In the dynamic analysis below, we will examine the impact of market size shocks on innovation. Consider an idiosyncratic market size shock to product line j, denoted ε_j in the context our our model.[17] Equilibrium output and employment in line j are then shocked by the amounts $\frac{A_j \varepsilon_j y}{\gamma w}$ and $\frac{\varepsilon_j y}{\gamma w}$ respectively. Hence, holding innovation fixed, there will be a positive impact of ε_j on firm size in the short run, and this impact will be smaller for firms to the left of the threshold as these firms will not want to cross the threshold and bear the extra regulatory cost.

What is the effect of the impact of the shock on firm-level innovation? Equation (4) is modified by having the shock factor $(\varepsilon_j)^{\frac{1}{\eta-1}}$ pre-multiplying each term of the equation. The

[17]A common shock to all firms can be modeled as an increase in y. This will not have a differential effect on innovation in firms of different size, as all variables in our model are expressed in units of final output.

Figure 4: Aggregate economy-wide innovation as a function of the intensity of regulation



Notes: We simulate the amount of aggregate innovation in different economies relative to an unregulated benchmark economy as the intensity of regulation changes as indicated by the magnitude of the implicit tax (τ). For example, if $\tau = 0.02$, aggregate innovation is 0.96 relative to the benchmark, i.e. 4% lower. Parameter values are the same in regulated and regulated economies (as in notes to Figure 1) except we vary the value of τ.

marginal effect of the demand shock ($\frac{\partial z_j}{\partial \varepsilon_j}$) is derived by differentiating equation (4) with respect to ε_j. The impact of the shock on innovation intensity will be largest for small firms far below the regulatory threshold. The second biggest effect will be on innovation in large firms well to the right of the threshold. And the smallest effect of the demand shock will be on firms just below the threshold.

In the data, the relationship between a shock and absolute innovation is complicated by the fact the marginal effect will also scale with size. We look at total innovation $Z = nz(n)$, so even abstracting away from the direct effect of the demand shock on size, the marginal effect of the demand shock on total firm innovation will be $\frac{\partial Z_j}{\partial \varepsilon_j} = n\frac{\partial z_j}{\partial \varepsilon_j}$. Thus larger firms will tend to respond more to the shock than smaller firms. However, even controlling for firm size (as we will do in the empirical work) and so concentrating on the marginal effect of the shock on innovation intensity, the model predicts that the effect of a market size shock on innovation should be significantly lower for firms just to the left of the threshold as $\frac{\partial z_j}{\partial \varepsilon_j}$ is smallest for these firms from equation (4).

Finally, the shock will affect the firm size distribution. If the shock is transitory, a shocked firm will grow larger for a short period of time before the economy will return to the initial steady state distribution. A permanent idiosyncratic shock will translate into a permanent change to the overall steady state size distribution. The dynamic empirical design is not well suited to analyzing the impact on the steady state firm size distribution as the shock is defined only for incumbents. Hence, we focus on entry effects in the equilibrium calibration.

3 Empirics

3.1 Data

Our main data comes from the French tax authorities, which consistently collect information on the balance sheets of all French firms on a yearly basis from 1994 to 2007 ("FICUS"). We restrict attention to non-government businesses and take patenting information from Lequien et al. (2017). This matches the PATSTAT (Spring 2016) database to FICUS using an algorithm, which matches the name of the affiliate - the holder of the IP rights - on the patent front page to the firm whose name and address is closest to that of the patent holder. The accuracy of the algorithm is worse for firms that are below 10 employees so we focus on firms with more than 10 employees. Since we are interested in the effects of a regulation that affects firms as they pass the 50 employees threshold, we further restrict attention to firms with between 10 and 100 workers in 1994 (or the first year those firms appear in the data).[18] More details about the data source are given in Appendix B.

Our main sample consists of 182,348 distinct firms and 1.66 million observations. We report some basic descriptive statistics in Table 1. We can see that on average, firms file 0.009 priority patents per year and, conditional on being an innovator, 0.28 per year. As is well known, the distribution of innovation is highly skewed, with a small number of firms owning a large share of the patents in our sample. However, since we do not include the largest French firms in our data, the skewness is less pronounced.

3.2 Nonparametric evidence: Static Analysis

Figure 5 shows, for each employment size bin, the fraction of firms within that bin with at least one patent (see also Panel A of Table 1). We see an almost linear relationship between firm size and the fraction of innovative firms. That larger firms are more likely to patent is in line with the analysis in Akcigit and Kerr (2018). The prediction of a linear relationship between firm size and innovation is consistent with our equation (4).

[18] We show robustness of the results to changing this bandwidth (see in particular Table D2 in Appendix D).

Table 1: Descriptive statistics

Panel A: All firms	Mean	p25	p50	p75	p90	p99
Employment	29	12	19	35	56	151
Sales	5,434	958	2,032	4,756	10,632	45,224
Patents	0.0090	0	0	0	0	0
Innovative	0.031	0	0	0	0	1
Manufacturing	0.25	0	0	1	1	1

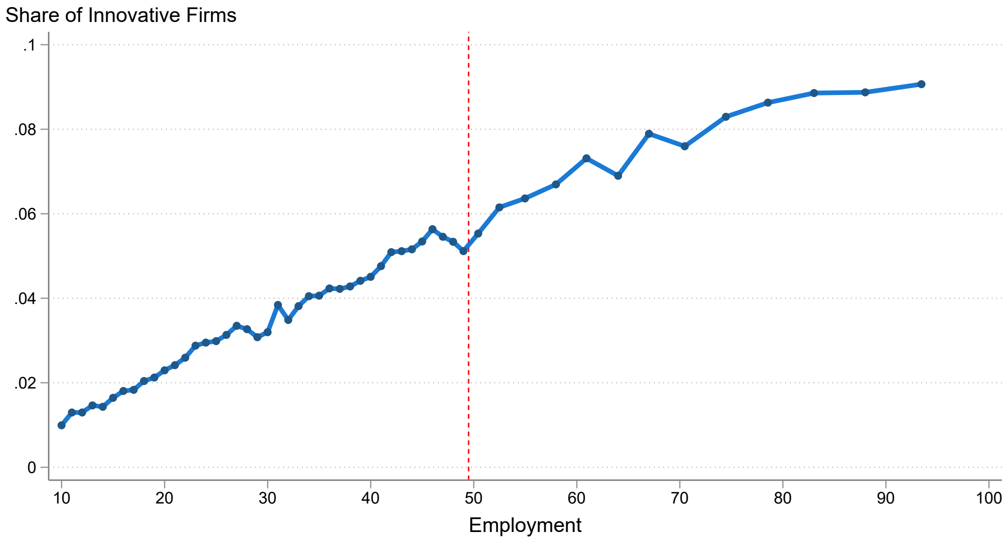
Panel B: Subset of innovative firms	Mean	p25	p50	p75	p90	p99
Employment	52	20	35	62	98	307
Sales	11,795	2,500	5,208	10,492	21,326	115,145
Patents	0.283	0	0	0	1	4
Manufacturing	0.68	0	1	1	1	1

Notes: These are descriptive statistics on our data. Panel A is all firms and Panel B conditions on firms who filed for a patent at least once over the 1994 to 2007 period ("Innovative" firms). We restrict to firms who have between 10 to 100 employees in 1994 (or the the first year they enter the sample). There are 182,347 firms and 1,658,762 observations in Panel A and 4,084 firms and 51,192 observations in Panel B.

For firms just below the 50 employee threshold, the share of innovative firms suddenly decreases in an innovation valley. This is what the model predicts. It is also noteworthy that the slope of the innovation-size relationship is flatter for larger firms to the right of the threshold than for smaller firms below the threshold. This again is consistent with our theoretical predictions. Note that in the theory, the ratio between the slopes of the innovation-size relationship between a large and a small firm, varies with the tax (τ) and with the concavity of the R&D cost function (η). We will exploit this variation to help recover the tax parameter later in this section.

The innovation outcome measure is taken over the whole sample period from 1994 to 2007, but the same patterns emerge if we consider alternative definitions of an innovative firm (see Appendix Figure D2). The predictions over the size distribution also broadly match up to the data, but since these are relatively well known we relegate discussion to Appendix D.

Figure 5: Share of innovative firms at each employment level



Notes: Share of innovative firms (i.e. with at least one priority patent) plotted against their employment. All observations are pooled together. Employment bins have been aggregated so as to include at least 10,000 firms. The sample is based on all firms with initial employment between 10 and 100 (182,347 firms and 1,658,762 observations, see Panel A of Table 1).

3.3 Dynamic analysis

3.3.1 Estimation equation

We now turn to our parametric investigation of how firms respond to market size shocks. More specifically, we estimate the regression:

$$\begin{aligned}\tilde{\Delta} Y_{i,t} &= \beta L^{\star}_{i,t-2} + \gamma[\Delta S_{i,t-2} \times \mathcal{P}(log(L_{i,t-2}))] + \delta[\Delta S_{i,t-2} \times L^{\star}_{i,t-2}] \\ &\quad + \phi \mathcal{P}(log(L_{i,t-2})) + \psi_{s(i,t)} + \tau_t + \varepsilon_{i,t} \qquad (6)\end{aligned}$$

where: $Y_{i,t}$ is a measure of innovation and $L_{i,t}$ a measure of employment; $L^{\star}_{i,t}$ is a binary variable that takes value 1 if firm i is close to, but below, the regulatory threshold at time t. Our baseline measure of $L^{\star}_{i,t}$ is a dummy for a firm having employment between 45 and 49 employees. $\Delta S_{i,t-2}$ is an exogenous demand shock to market size that should trigger an increase of innovation in a wide class of models (including our own); $\psi_{s(i,t)}$ is a set of industry dummies[19] and τ_t is a set of time dummies ($s(i,t)$ denotes the main sector of activity of firm i at time t), $\mathcal{P}(log(L_{i,t-2}))$ is a polynomial in $log(L_{i,t-2})$ and $\varepsilon_{i,t}$ is an error term. In our baseline empirical model we use a two year lag of the shock since there is likely to be some delay between the demand shock, the increase in research effort and the filing of a patent application. But we

[19]We allow these to change over time if a firm switches industries, but nothing really changes if we keep the industry definitions constant across time.

show robustness to other lag lengths. Finally, for the dependent variable we use growth rates of Y defined as:[20]

$$\tilde{\Delta}Y_{i,t} = \begin{cases} \frac{Y_t - Y_{t-1}}{Y_t + Y_{t-1}} & \text{if } Y_t + Y_{t-1} > 0 \\ 0 & \text{otherwise} \end{cases}$$

The key coefficient in the regression is δ, which we expect to be negative. Larger firms will likely respond more to a given shock, but this relationship should break down for the firms just to the left of the threshold as such firms are more reluctant to cross the threshold in response to a positive market size expansion.

3.3.2 Market Size Shocks

To construct the innovation shifters $\Delta S_{i,t-2}$, we rely on international trade data to build export demand shocks following Hummels et al. (2014) and Mayer et al. (2016). In short, we look at how foreign demand for a given product changes over time by measuring the change in imports from all countries worldwide (except France). We then build a product/destination portfolio for each French firm i, and weight the foreign demands for each product by the relative importance of that product for firm i. More specifically, firm i's export demand shock at date t is defined as:

$$\Delta S_{i,t} = \sum_{s,c \in \Omega(i,t_0)} \omega_{i,s,c,t_0} \tilde{\Delta} I_{s,c,t}, \tag{7}$$

where: $\Omega(i, t_0)$ is the set of products and destinations associated with positive export quantities by firm i in the first year t_0 in which we observe that firm in the customs data[21] and ω_{i,s,c,t_0} is the relative importance of product s and country destination c for firm i at t_0, defined as firm i's exports of product s to country c divided by total exports of firm i in that year. $I_{s,c,t}$ is country c's demand for product s, defined as the sum of its imports of product s from all countries except France. The basic idea behind the shock design is simply that a firm that was exporting, for example, many cars to China in 2000, would have benefited disproportionately from the boom in Chinese consumption of cars at the start of the twenty-first century.

We fix the weights at the firm level taking initial period t_0 as the reference. This is done in order to exclude any variation in the portfolio of products and countries that could be endogenous. Our shock is therefore similar to a "Bartik"-type shift-share instrument. There

[20] This is essentially the same as in Davis and Haltiwanger (1992) for employment dynamics except that we set the variable equal to zero when a firm does not patent for two periods. Results are robust to considering other types of growth rates (e.g. see the last three columns of Table D2 in Appendix D).

[21] French customs data are available from 1994. So we use 1994 as the initial year, except for firms who enter after 1994. For the entrants we use the initial year they enter the sample.

is an important recent literature (e.g. Borusyak et al., 2018, Goldsmith-Pinkham et al., 2020, Adao et al., 2019) which discusses inference and estimations with such variable. Importantly for us, the sum of exposure weights w_{f,j,s,t_0} across (s, j)'s is different from 1 and varies across firms. We follow Borusyak et al. (2018) who argue that in such an "incomplete shift-share" case with panel data, it is important to control for this sum and allow the coefficient to change with time.[22]

3.3.3 Testing the main prediction

To estimate equation (6), we need to make some further restrictions in our use of the dataset. First, note that the market size shock ΔS is only defined for exporting firms, that is, firms that appear at least once in the customs data from 1994 to 2007. Second, in order to increase the accuracy of our shock measure, we restrict attention to the manufacturing sector. Not only is a large fraction of patenting activity located in manufacturing, but these firms are also more likely to take part in the production of the goods they export (see Mayer et al., 2016). Our main regression sample is therefore composed of 24,081 firms and 189,727 observations.

Table 2 presents the results from estimating equation (6), i.e. from regressing the growth rate of firm patents on the lagged market size shock. Column (1) shows that firms facing a positive exogenous export shock are significantly more likely to increase their innovative activity. The coefficient implies that a 10% increase in market size increases patents by about 3%. Column (2) includes a control for the lagged level of log(employment) and also its interaction with the shock. The interaction coefficient is positive and significant, indicating that there is a general tendency for larger firms to respond more to the shock than smaller firms. Although it is not of direct interest, this is what we should expect given our discussion in 2.5. Column (3) generalizes this specification by adding in a quadratic term in lagged employment and its interaction with the shock.

[22] We have conducted many more extensive diagnostic tests showing the validity of this source of exogenous variation to market size. Importantly for us, Borusyak et al. (2018) underline two assumptions to address the validity of a shift-share instrument: quasi-randomness of shock assignment and high number of uncorrelated shocks. The first assumption is more likely to be true in our setting thanks to the inclusion of industry fixed effects in our regressions. The assumption is essentially that within industry, the expected value of ΔI is the same for all firms conditional on the country-product-level unobservables. The second assumption is warranted by the fact that we consider a very large number of shocks across many countries and products. In one robustness test, we follow the recommendations of Borusyak et al. (2018) and check that our main results are robust to using alternative shocks in which ΔI has been residualized on different combinations of year, country, product fixed effects. Moreover, note that our panel data structure allows us to include a firm fixed effect as an additional robustness check which further controls for potential correlations between permanent firm characteristics and future realizations of the shocks. See Aghion et al. (2018a) for more diagnostics.

Table 2: Main regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Shock_{t-2} \times L^{\star}_{t-2}$				-8.530**	-10.738***	-9.772***	-12.580**	-7.456**	-11.729**
				(3.122)	(3.419)	(3.258)	(4.868)	(3.368)	(4.202)
$L^{\star}_{t-2}$				-0.002	0.026	0.026	0.105	0.052	0.087
				(0.099)	(0.098)	(0.098)	(0.150)	(0.059)	(0.084)
$Shock_{t-2}$	2.316**	-9.331**	11.472	2.858***	-10.213**	9.392	-10.573*	-6.124**	-10.258**
	(0.919)	(4.221)	(10.268)	(0.888)	(4.281)	(10.236)	(5.123)	(2.488)	(4.024)
$log(L)_{t-2}$		-0.052	-0.030		-0.054	-0.029	-0.314	-0.028	-0.066**
		(0.037)	(0.133)		(0.037)	(0.135)	(0.187)	(0.019)	(0.032)
$Shock_{t-2} \times log(L)_{t-2}$		3.487**	-10.159		3.972***	-8.881	4.082**	2.277**	3.928***
		(1.340)	(7.947)		(1.387)	(7.893)	(1.696)	(0.875)	(1.314)
$log(L)^2_{t-2}$			-0.004			-0.004			
			(0.024)			(0.024)			
$Shock_{t-2} \times log(L)^2_{t-2}$			2.078			1.951			
			(1.378)			(1.371)			
$\Delta log(L)_{t-2}$									0.076
									(0.188)
Fixed Effects									
Sector×Year	✓	✓	✓	✓	✓	✓	✓	✓	✓
Firm							✓		
Number Obs.	189,727	189,727	189,727	189,727	189,727	189,727	189,727	398,399	172,323

Notes: This contains OLS estimates of equation (6) on the manufacturing firms in Panel A of Table 1 who have exported at some point 1994-2007. Dependent variable is the Davis and Haltiwanger (1992) growth rate in the number of priority patent applications between $t-1$ and t. Column 1 only considers the direct effect of the shock, taken at $t-2$, column 2 uses a linear interaction with $log(L)$ taken at $t-2$ and column 3 considers a quadratic interaction. Columns 4, 5 and 6 do the same as columns 1, 2 and 3 respectively but also includes an interaction with $L^{\star}$, a dummy variable for having an employment size between 45 and 49 employees at $t-2$. Column 7 replicates column 5 but adds firm fixed effects. Column 8 includes non-manufacturing firms and column 9 also controls for the growth in log(employment) at $t-2$. All models include a 2-digit NACE sector dummies interacted with year dummies. Estimation period is 2007-1997. Standard errors are clustered at the 2-digit NACE sector level. ***, ** and * indicate p-value below 0.01, 0.05 and 0.1 respectively.

Column (4) of Table 2 returns to the simpler specification of column (1) and includes a dummy for whether the firm' employment is just below the regulatory threshold in the 45-49 employees range (defined as $L^{\star}$) at $t-2$, and the interaction of this dummy with the shock. Our key coefficient is on this interaction term, and it is clearly negative and significant as our model implies. This is one of our key results: innovation in firms just below the regulatory threshold is significantly less likely to respond to positive demand opportunities than in firms further away from the threshold. Our interpretation is that when a firm gets close to the employment threshold, it faces a large "growth tax" due to the regulatory cost of becoming larger than 50 employees. Consequently, such a firm will be more reluctant to invest in innovation in response to this new demand opportunity. The firm might even simply cut its innovative activities altogether to avoid the risk of crossing the threshold. We depict the relationship between innovation and the shock in Figure 6. The figure plots the implied marginal effect of the market size shock on innovation for different firm sizes using the coefficients in column (5) of Table 2. We see that innovation in larger firms tends to respond more positively to the export

Figure 6: Marginal effect of a market size shock on innovation

Notes: marginal effect of a shock at different level of employment, based on the model in column 5 Table 2.

shock than in smaller firms, but at the regulatory threshold there is a sharp fall in the marginal effect of the demand shock, consistent with our model (e.g. see subsection 2.5).

It might be the case that the negative interaction between the threshold and the shock could be due to some omitted non-linearities. Hence in column (5) we also include lagged employment and its interaction with the shock (as in column (2)). These do have explanatory power, but our key interaction coefficient remains significant and negative and we treat this as our preferred specification. Column (6) adds a quadratic employment term and its interaction following column (3). Our key interaction remains significant and these additional non-linearities are insignificant.

3.3.4 Robustness of the dynamic empirical model

We have subjected our results to a large number of robustness tests, many of which are detailed in Appendix D. Column (7) of Table 2 shows the results from a tough robustness test where we include a full set of firm dummies. Given that the regression equation is already specified in first differences, this amounts to allowing firm-specific time trends. The key interaction between the market size shock and the threshold dummy remains significant. The data sample underlying Table 2 is limited to manufacturing firms. Column (8) also adds in non-manufacturing firms. The relationship remains negative, though with a smaller coefficient. This is likely to be due to the fact that patents are a much more noisy measure of innovation in non-manufacturing firms.

Does the number of patents grow more slowly for firms to the left of the threshold who experience a demand shock simply because their employment grows by less? Column (9) of 2 provides a crude test of this hypothesis by including the growth of employment on the right hand side of the regression. This variable is endogenous, of course, yet it is interesting to observe from a purely descriptive viewpoint that the interaction between the market size shock and the threshold remains significant. This suggests that it is patenting per worker, which is reacting negatively to the interaction between the shock and the threshold: our effect on patenting is not simply reflecting differential changes in firm size.

4 The aggregate effects of regulation on Innovation

So far, we have established that many of the *qualitative* predictions of our simple model are consistent with the data both from a non-parametric cross sectional analysis and a more challenging dynamic analysis of the response to shocks. In this section, we use the data, the structure of our theoretical model and some external calibration values to estimate the general equilibrium effects of the regulation on aggregate innovation and welfare. This clearly requires stronger assumptions as we are extrapolating well away from the threshold.

4.1 Quantitative Strategy

The details of our calculations are in C.1, but we sketch some of the important elements here. The threshold number of product lines, $\bar{n}$, can be calculated from the known regulatory employment threshold of 50, i.e. $\bar{n} = 50\omega\gamma$ (see equation (3)), so we have six unknown parameters: $(\eta, \omega, \gamma, \beta, \zeta, \tau)$. Since we only need the ratio β/ζ to calculate the aggregate innovation loss, we only need to quantify five parameters $(\eta, \omega, \gamma, \beta/\zeta, \tau)$. We use the existing literature to obtain two of them (η and γ) and the remaining three are chosen to match moments from the data as detailed in Table 3.

Table 3: Calibration values and moments

Parameter	Value	Name	Source
From the literature			
η	1.5	Concavity of Innovation cost function	Dechezlepretre et al. (2016)
γ	1.3	Productivity step size	Aghion et al. (2019a)
Using our data			
τ	0.025	Regulatory tax	Innovation-Firm size relationship $(\hat{\beta}_1, \hat{\beta}_2)$
β/ζ	1.65	Discount factor/scale parameter	Long-term growth of GDP
ω	0.25	Output adjusted wage	Changes in the firm size distribution

Concavity of the R&D cost function η In order to calibrate the concavity of the R&D cost function, η, we draw upon existing work that has estimated the innovation production function (the relationship between patents and R&D). Acemoglu et al. (2018) use a value of η = 2 based on Blundell et al. (2002). However, these estimates typically come from very large US firms (publicly listed companies from Compustat), so may exaggerate η, which is likely to be lower for the small and medium sized enterprises that are the focus of our study.[23] Indeed, the estimates of Dechezlepretre et al. (2016) which look at firms of similar sizes to the ones we use here, suggest a value of $\eta = 1.5$, using their Regression Discontinuity Design, which should produce cleaner causal estimates of the impact of R&D on innovation. This value is also consistent with some of the estimates in Crépon and Duguet (1997) on French firm panel data.

Regulatory tax τ To quantify the regulatory tax (τ), we estimate empirically the changing slope of the relationship between innovation and firm size from equation (4). Our theory implies that the ratio of the innovation-size slope for small firms (before the innovation "valley") to large firms (to the right of the regulatory threshold) should be equal to $(1-\tau)^{\frac{1}{\eta-1}}$. In other words, for any given value of η, a larger tax will mean a greater flattening of the positive relationship between innovation and firm size. Figure 5 shows this flattening very clearly and we recover this through a simple regression of patents on lagged size for firms under 45 employees and firms firms over 50 employees (to abstract from the innovation valley), allowing the coefficient on size to be different for these two size groups. Empirically, we average the number of patent applications filed by a firm over a five-year window for each possible value of employment L between 10 and 100. We then fit two different slopes for $L \in [10; 45]$ and $L \in [50; 100]$. We respectively denote $\hat{\beta}_1$ and $\hat{\beta}_2$ the OLS estimate of these two slopes.

We find $\hat{\beta}_1/1000 = 0.4031$ with a standard error $(\hat{\sigma_1}/1000)$ of 0.0206 and $\hat{\beta}_2/1000 = 0.3832$ with a standard error($\hat{\sigma_2}/1000$) of 0.0082. Hence, according to our model we have:

$$\frac{\hat{\beta}_1}{\hat{\beta}_2} = (1-\tau)^{\frac{1}{\eta-1}} = 0.951$$

Given the calibrated value of $\eta = 1.5$ this yields an estimate of $\tau = 0.025$, a regulatory tax of 2.5 percent. There are several ways to estimate this slope and we discuss the sensitivity to the choice of alternative empirical models extensively in Appendix D.4. Alternative models

[23] Labelling the estimated elasticity between patents and R&D as θ, $\eta = 1/\theta$. Since θ is likely larger for small firms (e.g. due to financial constraints) or in countries with less developed risk capital markets (e.g. France vs. the US) this implies a smaller η.

generate implicit taxes in the range of 1.1% to 4.9%, so we are effectively choosing a calibration value just below the midpoint of this range.

Step size γ The productivity step size γ following innovation is set to 1.3 using based on estimates in Aghion et al. (2019a). This is derived from various estimates of the average markup, which in our model is the reward from innovation.

Productivity adjusted wage rate ω A larger ω means a higher cost of labor and therefore a smaller mass of large firms. Therefore to set the value of ω, we use the empirical firm size distribution. In particular, we match the fall in the density of employment of smaller vs. larger firms to the left and right of the innovation valley. In our data there are about three time as many firms between 40 and 45 employees than between 50 and 55 and the value of ω that reproduces this gap is 0.25.

Scale parameter and discount factor β/ζ We calibrate β/ζ in order to match the measured value of g in the data that we take to be equal to the average growth of GDP in France over the period 1990-2019 (1.62%). In our model, growth g is defined as follow:

$$g = \exp\left((z_e + \sum_{i=1}^{N} \mu(i)z(i)i)\log(\gamma)\right) - 1.$$

This yields a value of β/ζ of 1.65.

4.2 Results

4.2.1 Baseline result:

From Figure 4 we can see that $\tau = 0.025$ implies a loss of aggregate innovation of about 5.4% percent compared to the no regulation benchmark. As discussed in the modeling section, the aggregate loss is driven by three major elements:

1. The decline in the incumbent innovation rate ($z(n)$) for a given firm size. For any given size distribution of firms, the regulation reduces innovation rates for firms above the threshold and just to the left of the threshold.
2. The change in the size distribution μ. Since the regulation pushes the size distribution to the left and smaller firms do less innovation, this reduces aggregate innovation.

3. The decline in the innovation rate by entrants z_e.

Recall that we have denoted $\mathcal{Z}(\tau) = \sum_{i=1}^{\infty} \mu(i)z(i)i + z_e$ total innovation in the economy when the regulation tax is set to τ and the value of other variables are taken from Table 3. Analogously to a shift-share decomposition analysis we have:

$$\begin{aligned}
\mathcal{Z}(\tau) - \mathcal{Z}(0) \;=\; & \sum_{n>0} \left(Z(n,\tau) - Z(n,0)\right) \mu(n,0) \\
& + \sum_{n>0} \left(\mu(n,\tau) - \mu(n,0)\right) Z(n,0) \\
& + \sum_{n>0} \left(\mu(n,\tau) - \mu(n,0)\right) \left(Z(n,\tau) - Z(n,0)\right) \\
& + z_e(\tau) - z_e(0),
\end{aligned} \tag{8}$$

where $\mu(n,\tau)$ and $Z(n,\tau)$ are the share of firm of size n where the economy has a regulation tax of τ and their total innovation respectively. The first term in the right hand side of equation (8) is the innovation intensity (evaluated at the size distribution in the unregulated economy) and the second term is the effect on size (evaluated at a firm's innovation intensity rate in the unregulated economy). The third term is the interaction effect between the first two terms and the final term is the effect on entrants (since an entrant must innovate by definition to displace an incumbent).

Dividing equation (8) by $\mathcal{Z}(0)$, we can have an approximation of where the 5.4% loss of aggregate innovation comes from. We find that most (80%) of the effect comes from the change in the innovation intensity (the first term in the right hand side of the previous equation). The covariance and entry terms (third and last terms) account for roughly 10% each, while the change in the size distribution has almost no effect. The virtual absence of any effect of the size distribution is because the value of the tax is relatively small.

4.2.2 Robustness of the Aggregate Calculations

We now explore how the 5.4% loss in innovation is affected when we consider variations in the parameters from Table 3. In Table 4, we consider the effect of changes in η, γ, ω, τ and β/ζ. With respect to η, we consider the range interval $\eta \in [1.3, 2]$ to reflect the variety of values found in the literature (see above). With respect to γ, we explore values from 1.2 to 1.5. A value of 1.5 corresponds to a labor share of 66% in our model.[24] Regarding ω, and β/ζ, we consider a relative change of 15% (upward and downward) .

Table 4: Sensitivity analysis

Robustness	**Loss in total innovation**
Baseline	5.42%
$\gamma = 1.2$	5.34%
$\gamma = 1.50$	5.45%
$\eta = 2$	2.76%
$\eta = 1.3$	8.80%
$\omega = 0.22$	5.35%
$\omega = 0.29$	5.45%
$\beta/\zeta = 1.40$	5.42%
$\beta/\zeta = 1.90$	5.42%
τ	
Percentile 75^{th} ($\tau = 0.043$)	9.75%
Percentile 25^{th} ($\tau = 0.007$)	1.40%

Notes: baseline uses parameter values: ($\eta = 1.5$, $\gamma = 1.3$, $\tau = 0.025$, $\beta/\zeta = 1.65$ and $\omega = 0.25$), see Table 3. In the robustness where γ, η, ω or β/ζ are changed, we keep τ as in the baseline. The last two lines report the 25^{th} and 75^{th} percentile for the loss of innovation in a sample computed from 100,000 independent draws of τ from two normal distribution. The corresponding value of τ and β/ζ are computed as an average for each percentile.

Given that τ has been calculated using estimates of the slopes of the innovation / size cross-section, we use our estimates of β_1 and β_2 to derive confidence intervals for τ. Specifically, we draw 100,000 values of β_1 and β_2 from two independent normal distribution $\mathcal{N}(\hat{\beta}_1, \hat{\sigma_1})$ and $\mathcal{N}(\hat{\beta}_2, \hat{\sigma_2})$, where $\hat{\beta}_i$ and $\hat{\sigma_i}$ respectively designate the point estimates and corresponding standard errors. For each of these 100,000 draws, we compute a value for τ and infer the loss in total innovation and welfare by running the model.

The results from this exercise can be found in Table 4.

As we would expect, the most important parameter is the regulatory tax, τ. From the values of β_1 and β_2, the loss is 9.8% for the 75^{th} percentile of the distribution and 1.4% for at the 25^{th} percentile (the median is the same as the baseline: 5.4%). Interestingly, η also matters: as the parameter moves from 1.3 to 2, the aggregate innovation losses falls from 8.8% to 2.8%.

[24] In a wide class of models the ratio of price to marginal cost (the markup) is equal to the output elasticity with respect to a variable factor of production divided by the variables factor's share of revenue (e.g. De Loecker et al., 2020; Hall, 1988). Since labor is the only factor in our model, the markup is simply the reciprocal of the labor share. Aghion et al. (2019a) use a a US labor share of GDP of 77% to obtain $\gamma = 1.3$. The French labor share after 1995 is more like 65% (see e.g. Cette et al., 2019), suggesting $\gamma = 1.5$. These values encompass most of the other estimates of the aggregate markup using other methods.

This is because changing η determines the elasticity of innovation with respect to R&D: as η increases, the impact of R&D on innovation decreases. Since the impact of the tax comes from reducing the incentive to do R&D to grow, if R&D has little effect on growth there will be little impact of the tax. Hence, increasing η makes total innovation less sensitive to changes in τ.

By contrast, the loss in total innovation is only modestly affected by changes in γ, ω and β/ζ. This is because the tax elasticity of z ($\frac{dz}{d(1-\tau)}\frac{1-\tau}{z}$) only depends upon η, not on ω, γ or β/ζ. From equation (4), we see that the elasticity of innovation with respect to the regulatory tax is $\frac{1}{\eta-1}$ for large firms. Hence, changing the values of ω, γ and β/ζ only affects total innovation loss through their effects on the firm size distribution and on entry, which we know from the previous subsection plays a relatively minor quantitative role.[25]

4.3 Welfare

Innovation increases growth which is a benefit to welfare, but it must also be paid for by diverting current consumption into R&D investments. In Schumpeterian growth models, the impact of a reduction in innovation on welfare is theoretically ambiguous. Although positive knowledge externalities generate the traditional underinvestment in R&D, the business stealing effect can generate too much investment. Which dominates in our setting? Using the utility of the representative agent in equation (1), C_t is determined by the final good market clearing condition which states that each unit of final good that is produced should be used either for consumption C_t or $R\&D$. Recall that to produce an innovation intensity of $Z = nz$, a firm must spend $\zeta n z^\eta$ units of final good. We therefore have the following identity:

$$Y_t = C_t + \sum_{i\geq 1} \zeta \mu(i) i z(i)^\eta Y_t,$$

i.e. we take away R&D expenditures (there are $\mu(i)$ firms of size i) from the final good Y_t, and the residual is consumed. Denoting aggregate R&D $R \equiv \sum_{i\geq 1} \zeta\mu(i)iz(i)^\eta$ and plugging this into the utility function yields:

$$U = \sum_{t>0} \beta^t \log(Y_0(1+g)^t(1-R)),$$

which can be rewritten:

[25] For example, as already noted a higher ω reduces the relative numbers of large firms. Since there are more firms just to the left of the regulatory threshold (whose innovation is most affected by the regulation), this makes the marginal impact of the tax slightly larger.

$$U = \frac{\log(Y_0)}{1-\beta} + \frac{\log(1+g)\beta}{(1-\beta)^2} + \frac{\log(1-R)}{1-\beta}$$

Since growth is $g = \left(z_e + \sum_{i\geq 1} iz(i)\mu(i)\right)\log(\gamma)$ and using the definition of R, we can compute total utility for any value of Y_0 using vectors z and μ and the value of z_e.

We define $g(\tau)$, $R(\tau)$ and $Y_0(\tau)$ the values of g, R and Y_0 in an economy with a regulation level equal to τ. Let

$$\begin{aligned}\Delta U &\equiv U(\tau) - U(0)\\ &= \log\left(\frac{1+g(\tau)}{1+g(0)}\right)\frac{\beta}{(1-\beta)^2} + \log\left(\frac{1-R(\tau)}{1-R(0)}\right)\frac{1}{1-\beta} + \log\left(\frac{Y_0(\tau)}{Y_0(0)}\right)\frac{1}{1-\beta},\end{aligned}$$

denote the difference in utility between an economy with regulation τ and an economy without regulation at the steady-state. The corresponding difference in terms of consumption equivalent is given by $\exp\left((1-\beta)\Delta U\right)$. Initial production Y_0 is equal to initial quality times the amount of labor used in production. In our baseline model, the whole labor force is employed in production with and without the regulation, as R&D does not require labor.[26] Hence, abstracting from initial quality, the effect of the regulation on welfare is governed by the first two terms in the above equation.

The first term is negative since $g(\tau) < g(0)$ due to lower innovation, hence a welfare loss from introducing the regulation. The second term is positive $(R(\tau) < R(0))$: the corresponding welfare gain stems from the fact that spending less on R&D leaves more output for consumption. Thus the overall effect of the regulation on welfare is *a priori* ambiguous.

Yet, using our baseline parameter values from Table 3 and a standard value of $\beta = 0.96$, we can compute the difference in welfare (abstracting from initial quality) in terms of consumption equivalent. In our baseline regulated economy, welfare is 2.2% lower than in the unregulated economy. Table D4 in Appendix D.4 shows the welfare losses under the various alternative assumptions on the calibration values.[27]

[26]This is no longer true if labor is used in production and in R&D (see next section). Then the tax regulation will affect Y_0 even controlling for initial quality as it will affect the fraction of labor used in production.

[27]Measuring welfare requires a separate estimation of β and ζ. The measure of welfare is obviously sensitive to the choice of β. Specifically, welfare loss will increase as β is closer to 1 as agent gives more weight to future consumption and therefore care more about growth. When $\beta = 0.94$, welfare loss amount to 1.39% while when $\beta = 0.98$, welfare loss is 4.4% (see Table D4 in Appendix D.4).

4.4 Summary on the Aggregate innovation effects of regulation

The effects of regulation on aggregate innovation appear non-trivial. The losses are around 5.4% in our baseline estimates and vary between 1.4% and 9.8% when we examine a wide range of different values for the parameters. Four-fifths of the losses come from a lower amount of innovation across all affected firms, with the residual fifth accounted for by lower entry and a leftwards shift of the firm size distribution. Our baseline results find a (lower bound) fall in welfare of 2.2% from these dynamic losses.

5 Extensions and the Nature of Innovation

Our baseline model focuses on the impact of regulation on the *rate* of innovation. But there are various ways in which regulations may affect the nature of innovation. In this section, we first consider an extension of our model which allows firms to invest simultaneously in two types of innovation: incremental or radical. After developing the theory we implement this empirically using two proxies for how radical a patent is: (i) a traditional future citations measure and (ii) a more novel machine learning algorithm based on the full text of the patent. Secondly, we also use textual patent analysis to measure automation as one response to the regulation may be to invest in labor saving innovations. Finally, we extend our analysis to allow for infinitely lived owners and for R&D as scientists.

5.1 Radical versus incremental innovation

Although regulation discourages overall innovation, it may also alter the *type* of innovation. A firm just below the threshold has a reduced incentive to innovate, but it might be that if she does innovate she will "swing for the fence" by investing in radical innovation. Minor, incremental innovations that just push the firm over the threshold will be strongly discouraged by the regulation. We now formalize this intuition and then test whether it has any relevance in the data.

5.1.1 Theory

In our baseline model, firms could only increase their number of product lines by one line in each period. In this extension, we assume that firms can now choose between: (i) Investing in an incremental innovation which augments the firm's size by one additional product line and

(ii) Investing in more radical innovation which is more costly but augments the firm's size by $k > 1$ product lines. We now have four cases depending on the value for n:

1. $n < \bar{n} - k$ in which case the firm is never taxed in period 2.

2. $n < \bar{n}$ and $n \geq \bar{n} - k$ in which case the firm is taxed in period 2 only if it successfully innovated with a radical innovation.

3. $n = \bar{n} - 1$ in which case the firm is taxed in period 2 if it innovates, regardless of the type of innovation.

4. $n \leq \bar{n}$ in which case the firm is taxed in period 1 and 2 (except if the firm is at $\bar{n} + 1$ but this won't affect the firm's decision)

The firm therefore chooses z and u so as to maximize:

$$
\begin{aligned}
& n\pi(n) + \beta n z(n) \left((n+1)\pi(n+1) - n\pi(n) \right) + \beta n u(n) \left((n+k)\pi(n+k) - n\pi(n) \right) \\
& + \beta n x \left((n-1)\pi(n-1) - n\pi(n) \right) - n\zeta \left(z(n) + u(n) \right)^{\eta} - n\alpha u(n)^{\eta},
\end{aligned}
\tag{9}
$$

where α denotes the additional cost of radical innovation.[28]

The steady-state firm size distribution is computed in exactly the same way as in the baseline model, except that the flow equation needs to be adjusted to account for radical innovation:

$$
n\mu(n) \left(u(n) + z(n) + x \right) = \mu(n-1)z(n-1)(n-1) + \mu(n+1)x(n+1) + \mu(n-k)(n-k)u(n-k),
$$

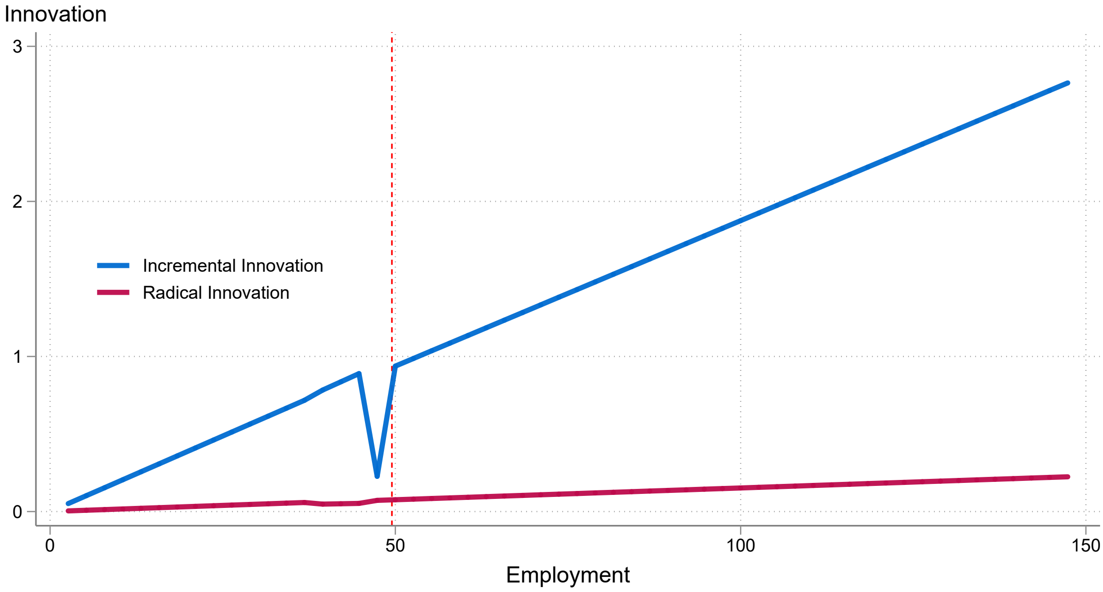
with $u(n-k)$ implicitly set to 0 if $n < k$.

We solve the model numerically using the same methodology and parameter values as in the baseline case and assume that $k = 4$ (that is, a successful radical innovation corresponds to a jump of 4 lines). We plot the new size distribution of firms against no tax $\tau = 0$. Results are presented in Figure C1 (Appendix C.2.1) and are qualitatively similar as in the model without radical innovation.

Regarding innovation intensity, we show in Figure 7 how the levels of incremental and radical innovation vary with firm employment size with and without regulation. We also plot the share of radical innovation over total innovation against employment in Figure 8. What these figures suggest is that the discouraging effect of the regulation on innovation by firms close to the threshold, is close to zero for radical innovations.

[28] In Appendix C, we solve formally for u and z in case where η=2 (quadratic R&D cost function).

Figure 7: Firm Innovation by employment size for incremental and radical innovations.



Notes: This is total incremental innovation $z(n)n$ (blue solid line) and total radical innovation $u(n)n$ (red dashed line) for firms of n lines against employment in the extension where firms can choose between two types of innovations. We used the same parameter values as in Figure 1 and $k = 4$ and $\alpha =$.

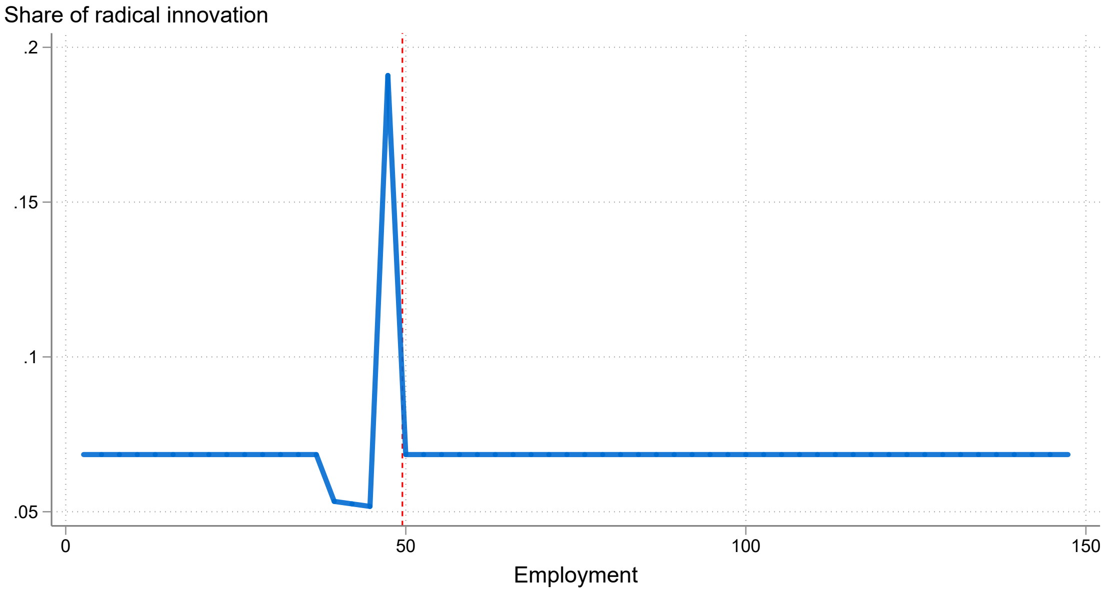
5.1.2 Evidence I: Citations

We first repeat the static analysis in Figure 9 using the quality of patents as the measure of innovation output. We measure quality using the number of future citations. For each patent within a technology class by cohort-year we determine whether the patent was in the top 10% most cited patents or in the bottom 90% (using future cites through to 2016). The two curves in Figure 5 correspond to the fractions of firms in each employment size bin respectively with patents in the top 10% cited and with patents in the bottom 90% cited. We clearly see that the drop in patenting just below the regulatory threshold is barely visible for patents in the top 10% cited. This is consistent with the idea that the regulation discourages low value innovation but not higher value innovation.[29] It is also clear from the figure that the innovation-size relationship is steeper for incremental innovation than for high value innovation. This is consistent with smaller firms accounting for a higher share of more radical innovation (e.g. Akcigit and Kerr, 2018, on US data and Manso et al. (2019)).

Next, we repeat our preferred dynamic specification of column (5) of Table 2, but now distinguish patents of different value using their future citations. Table 5 does this for patents

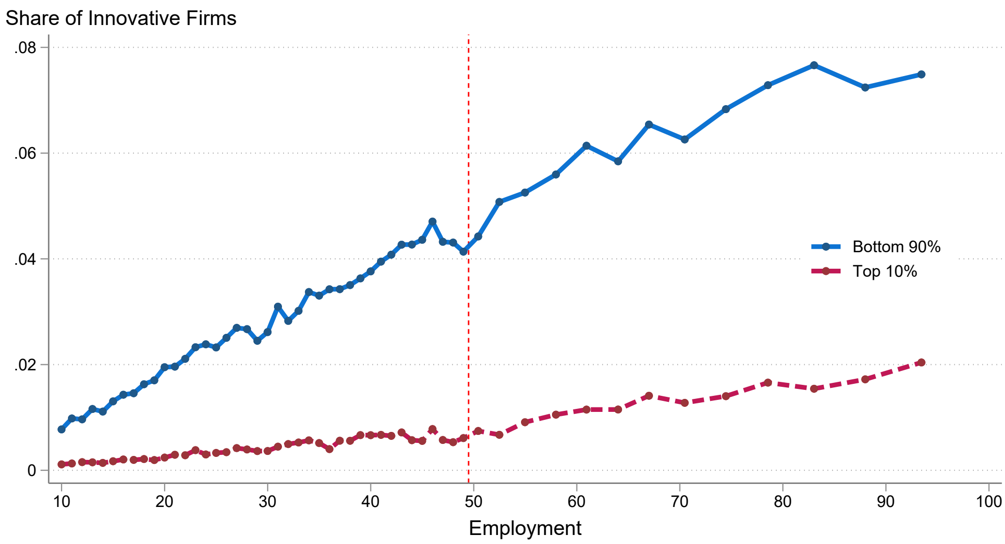
[29] As for Figure 5, Figure 9 considers the innovation outcome over the whole period of observations. Variants around this can be found in Figure D3 in the Online Appendix D.

Figure 8: Share of radical innovation in total innovation by firm employment size



Notes: This is the share of radical innovation over total innovation $u(n)/(z(n) + u(n))$ for firm with n product line against employment. We used the same parameter values as in Figure 7.

Figure 9: Share of innovative firms at each employment level and quality of innovation



Notes: Share of firms with at least one priority patent in the top 10% most cited (dashed line) and the share of firms with at least one priority patent among the bottom 90% most cited in the year (solid line). All observations are pooled together. Employment bins have been aggregated so as to include at least 10,000 firms. The sample is based on all firms with initial employment between 10 and 100 (82,347 firms and 1,658,762 observations, see Panel A of Table 1).

in the top 10%, 15% and 25% of the citation distribution in the first three columns and the patents in the complementary sets in the last three columns (i.e. the bottom 75%, 85% and 90% of the citation distribution). We clearly see that the negative effect of regulation on innovation is only significant for low quality patents in columns (4), (5) and (6). There are no such significant effects for patents in the top decile or quartile of the patent quality distribution (the coefficient on the interaction is even positive in column (2)).[30]

Table 5: Regression results for different levels of the quality of innovation

Quality	Top 10% (1)	Top 15% (2)	Top 25% (3)	Bottom 75% (4)	Bottom 85% (5)	Bottom 90% (6)
$Shock_{t-2} \times L^{\star}_{t-2}$	-0.365	1.471	-1.527	-7.937*	-10.532**	-10.695**
	(1.731)	(1.908)	(2.052)	(4.544)	(4.061)	(3.809)
$L^{\star}_{t-2}$	-0.052	-0.002	-0.046	0.103	0.017	0.028
	(0.035)	(0.068)	(0.066)	(0.108)	(0.098)	(0.103)
$Shock_{t-2}$	-3.631	-5.042*	-11.890***	-1.846	-7.218	-7.152
	(2.277)	(2.911)	(3.672)	(5.399)	(5.113)	(4.653)
$log(L)_{t-2}$	0.015	-0.011	-0.045	-0.001	-0.047	-0.063*
	(0.016)	(0.024)	(0.031)	(0.020)	(0.028)	(0.031)
$Shock_{t-2} \times log(L)_{t-2}$	1.148	1.558	3.705***	1.406	2.944	2.905*
	(0.714)	(0.929)	(1.183)	(1.950)	(1.771)	(1.616)
Fixed Effects						
Sector×Year	✓	✓	✓	✓	✓	✓
Number Obs.	189,727	189,727	189,727	189,727	189,727	189,727

Notes: estimation results of the same model as in column 5 of Table 2. The dependent variable is the Davis and Haltiwanger (1992) growth rate in the number of priority patent applications between $t-1$ and t, restricting to the top 10% most cited in the year (column 1), the top 15% most cited in the year (column 2), the top 25% most cited in the year (column 3), the bottom 85% most cited in the year (column 4), the bottom 75% most cited in the year (column 5) and the bottom 90% most cited in the year (column 6). All models include a 2-digit NACE sector interacted with year fixed effects. Estimation period: 1997-2007. Standard errors are clustered at the 2-digit NACE sector level. ***, ** and * indicate p-value below 0.01, 0.05 and 0.1 respectively.

To visualize these results, we plot the marginal effect of the demand shock on innovation by the level of firm employment in Figure 10. The dotted grey line is the marginal effect of the shock on patents in the bottom 90% of the quality distribution based on column (6) of Table 5. Overall, the impact of the shock is positive and larger for bigger firms. However, when we approach the regulatory threshold at 50, this relationship breaks down and the marginal effect of the shock falls precipitously (and actually becomes negative). The black solid line plots the marginal effect of the demand shock on high quality patents in the top decile of the citation distribution from column (1) of Table 5. This line is also positive for almost all firms and rises with firm size. By contrast, with low value patents, there is no evidence of any sharp downturn

[30] We show the diminishing effect of the shock around the threshold for many other quantiles of the patent value distribution in five percentile intervals in Figure D4. This shows a clearly declining pattern.

Figure 10: Total marginal effect of a shock

Notes: marginal effect of a shock at different level of employment, based on the model in column 1 and 6 of Table 5. Marginal effect is calculated on top 10% and bottom 90% most cited patents.

just below the regulatory threshold. [31]

In short, there seems to be evidence that the chilling effect of regulation on innovation is not an issue for high value patents and is instead confined to lower value patents, which is broadly consistent with the generalization of the model we developed for two types of R&D.

5.1.3 Evidence II: Patent text measures of novelty

We construct an alternative measure of radical innovation that is made to reflect the level of novelty of a patent using the text describing the patent (in the abstract and main body). We follow Kelly et al. (2018) who build an index of novelty by looking at how much the text of a given patent differs from the current state of knowledge in the technological classes using machine learning text-to-data techniques. This measure has been shown to capture features missed by citation-based indicators (see Bergeaud et al., 2017 for a review). For example, using many detailed industry case studies, the novelty measure has been shown to better reflect breakthrough technologies than citations (or other originality measures).

To implement this method we exploit the work of Google Patent (GP) who recently released a quantitative description of every patent that they describe as "embeddings" (see Srebrovic (2019) for details). GP embeddings use artificial intelligence analysis of text to summarize the

[31]The stronger relationship between demand growth and incremental (rather than radical) innovation is consistent with the earlier cross sectional Figure 9 and also Manso et al. (2019).

most important features of the patent text into a vector of 64 numbers bounded between -1 and 1. We can then calculate the "distance" between any pair of patents by simply taking the dot product between the two vectors. Full details are provided in Appendix D.5, but the basic idea is that we calculate novelty by computing the distance between a patent and a reference point from past patents in the same technological field. A more novel patent will use words that are further away from the current state of the art as indexed by the typical descriptions of patents.

We replicate all the analyses of the previous subsection on citations using this new measure in Appendix D and find broadly similar results. Note that this is not because the two measures are almost identical: the correlation between the two measures (cites vs. novelty) is only 0.1. First, in D5, we show that the cross sectional patterns show no innovation valley or a falling the innovation-size gradient at 50 employees for novel patents (in fact the gradient, if anything, is steeper after 50), whereas the usual patterns emerge for non-novel patents. Second, we replicate Table 5 and split patents between the top 10% , 15% and 25% and bottom 90% , 85% and 75% based on their novelty score. Table D6 shows that the least novel (bottom 90%) patents have a significantly lower response rate to the exogenous demand shock whereas there is a small and insignificant response of the top 10% most novel.

Broadly, both citation and novelty based measures of radical patents are consistent with the extension to the model to allow for endogenous types of R&D. The finding that the main effects of regulations are on incremental innovation implies some reduction in the magnitude of the welfare losses. Firm-level regressions in online Table D5 show that the effect of a radical innovation on firm growth is two to three times larger than the effect of an incremental innovation, but both types of innovations have a significant and positive effect. Hence, since most patents are incremental and these do matter for growth, although the negative effects of the regulation are smaller in our extended model, they are far from eliminated. [32]

5.2 Labor-Saving Technology

There are many ways in which firms can respond to the regulation other than by reducing the pace of innovation. In addition to cutting back employment growth, Garicano et al. (2016) document how firms approaching the threshold also increase overtime, capital investment, out-

[32] The exact magnitude of this mitigation effect is harder to calculate in the extended model due to its complexity, but we do present some back-of-the-envelope evidence in 5.2. This suggests adjusting the growth impact of regulation should be changed by about a fifth, from 5% to 4%.

sourcing and the skill mix. These might mitigate some of the costs, but will not eliminate the regulatory tax, as these are imperfect substitutes for job growth. Yet another strategy may be to develop labor saving automation technologies, that will enable the firm to increase output with less labor inputs.

To address the challenge of determining the degree to which a patent is about automation we again use textual analysis. In particular, we draw on Mann and Püttmann (2018) who used a supervised machine learning technique to classify automation and non-automation patents. Since their work was on the USPTO which is only a subsample of our data, we train an algorithm based on their classification using the GP embedding vector discussed in the previous subsection and then extrapolate this *predicted* measure of automation for all our sample. With this measure in hand, we again replicate all the analyses of the previous subsections. Consistent with our expectations, we find that the regulation only affected non-automation patenting (full results are presented in Appendix D.5). For example, Table D7 shows that faced with a positive demand shock, firms were significantly less likely to innovate in non-automation patents (bottom quartile), but were more likely to respond with automation patents (top quartile). Finally, we draw on a measure of process innovation developed by Ashish Arora and Lee (2020), which are more likely to be labor saving. This generates similar qualitative results to automation patents.

5.3 From two-period to infinitely-lived owners

In our baseline model, although firms can live forever we simplified the analytical problem by assuming the owners of firms only live for two periods. In this subsection, we show that the qualitative predictions of the model carry over to a more complex problem where owners and firms can be infinitely lived (or at least until the firm exits the market).

Let $V(n)$ denotes the value of a firm with n product lines. Each firm consider the following problem:

$$V(n) = \max_{z\geq 0} \left\{ n\pi(n)y - \zeta z^{\eta} ny + \frac{1}{1+r}\mathbb{E}\left[V(n')\right] \right\},$$

where $\mathbb{E}\left[V(n')\right] = nzV(n+1) + nxV(n-1) + (1 - nx - nz)V(n)$. Using the Euler equation, and letting $\rho = 1/\beta - 1$ and $W(n) = \beta V(n)/y$, we can rewrite the above Bellman equation as:

$$\rho\frac{W(n)}{n} = \max_{z\geq 0} \left\{ \pi(n) - \zeta z^{\eta} + z\left(W(n+1) - W(n)\right) + x\left(W(n-1) - W(n)\right) \right\}. \qquad (10)$$

Unlike our baseline model (which built on Klette and Kortum, 2004), the fact that $\pi(n)$ varies with n implies that a linear solution is no longer possible and thus we no longer have a closed-form solution for W. Nevertheless, we can still solve the model numerically. Details and results are given in Appendix C.2.2.

The key results carry over in this more complex model. In particular, although total innovation increases with size there is a fall in innovation for firms to the left of regulatory threshold. Figure C2 shows that this valley is somewhat smoothed out compared to our baseline over a larger mass of firms. This is because more firms anticipate that as they grow they will eventually get near to 49 where the tax on innovation is particularly large. Similarly, looking at the firm size distribution, the bump in mass is spread further to the left of 49 employees than in the baseline model for the same reason. This may help explain why the valley and bumps are smoother in the data than they are in the simple model. Since these changes are minor in a qualitative and quantitative sense, we prefer to keep to our simpler model as a baseline.

5.4 Endogenizing Equilibrium Wages

In the baseline model, R&D is a "lab equipment" model where the equipment is bought on the world market, labor supply is fixed and the labor force is all employed as production workers. This means the labor share, ω, is constant and unaffected by the regulation. In this extension, we consider the case where R&D uses scientists as an input, which means that the labor share can change with regulation. Full details are in Appendix C.2.3, but we sketch the main results here.

Workers can choose to supply labor to the R&D sector or to the production sector. In this case the total employment of firm i is given by:[33]

$$L_i = \frac{n_i}{\omega\gamma} + \zeta n_i z_i^{\eta} \equiv L(n_i, z_i),$$

and therefore L_i depends directly upon current innovation, instead of only through past innovation as reflected in its size $(\frac{n_i}{\omega\gamma})$. The employment threshold $\bar{l}$ no longer corresponds to a single number of products, but rather to a set of pairs (z, n) such that:

$$z = \frac{1}{\zeta n}\left(\bar{l} - \frac{n}{\gamma\omega}\right)^{\frac{1}{\eta}},$$

whenever $n \leq \bar{n}$.

As employment directly depends upon the level of z, so does the profit per line which is now equal to:

[33] Here ζ is a labor cost.

$$\pi(n,z) = \frac{\gamma - 1}{\gamma}\left(1 - \mathbb{1}\left[L(n,z) \geq \bar{l}\right]\tau\right)$$

The firm's problem is otherwise the same, but again the model needs to be solved numerically. Appendix C.2.3 shows that the qualitative effects again go through in terms of the size distribution and the firm innovation-size relationships. However, an important additional result is that the regulation reduces the equilibrium wage: the greater the tax, the greater the fall in the wage. This will mitigate the shift to the left in the size distribution.

6 Conclusion

In this paper, we have developed a framework to analyze the impact of regulation on innovation. We applied this to France, where strong labor regulations affect firms who employ 50 or more workers. We showed both theoretically and empirically that the prospect of these regulatory costs discourages firms just below the threshold from innovating, where innovation is measured by the volume of patent applications. This relationship emerges both when looking non-parametrically at patent density around the threshold and in a parametric exercise where we examine the heterogeneous response of firms to exogenous market size shocks (from export markets). On average, firms innovate more when they experience a positive shock, but this relationship significantly weakens when a firm is just below the regulatory threshold. We then use moments from our data and the literature to calibrate the structural parameters in the model. For example, using estimates of the R&D cost function, we can back out the magnitude of the regulatory tax from the ratio between the slopes of the innovation-size relationship for large firms compared to small firms. Our baseline estimates imply an aggregate innovation (and therefore growth) loss of about 5.4% and a lower bound on the loss of welfare of about 2.2%.

This suggests larger welfare losses than existing analyses that take technology as exogenous. A caveat to this conclusion is that when we use information on citations we find that the labor regulation deters incremental innovation, but has little effect on more radical innovation. This is consistent with a generalization of the model which allows for simultaneous investment in two types of R&D, and slightly mitigates the welfare loss of the regulation.

The analysis in this paper can be extended in several directions. First, our focus in this paper was on the long-run steady state, but it is perhaps equally important to analyze the transitional dynamics triggered by policy changes, and to factor in adjustment costs. Second, the framework can be applied to many other countries and regulatory settings. Third, our analysis remained focused on the costs of the labor regulation. However, such a regulation may also bring benefits in the form of better insurance and deeper involvement of employees in the management of the firm, which in turn fosters trust between employers and employees. Future work should take such benefits into account to see if they are sufficient to overcome the costs we have identified here.

References

Abowd, John M. and Francis Kramarz, "The costs of hiring and separations," *Labour Economics*, 2003, *15* (5), 499–530.

Acemoglu, Daron and Joshua Linn, "Market Size and Innovation: Theory and Evidence from the Pharmaceutical industry," *Quarterly Journal of Economics*, August 2004, *119* (3), 1049–1090.

_ , Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr, "Innovation, reallocation, and growth," *American Economic Review*, 2018, *108* (11), 3450–91.

Acharya, Viral V, Ramin P Baghai, and Krishnamurthy V Subramanian, "Labor laws and innovation," *The Journal of Law and Economics*, 2013, *56* (4), 997–1037.

_ , _ , and _ , "Wrongful discharge laws and innovation," *The Review of Financial Studies*, 2013, *27* (1), 301–346.

Adao, Rodrigo, Michal Kolesár, and Eduardo Morales, "Shift-share designs: Theory and inference," *The Quarterly Journal of Economics*, 2019, *134* (4), 1949–2010.

Aghion, Philippe, Antonin Bergeaud, Matthieu Lequien, and Marc Melitz, "The Impact of Exports on Innovation: Theory and Evidence," NBER Working Paper 24600, National Bureau of Economic Research 2018.

_ , _ , Timo Boppart, Peter J Klenow, and Huiyu Li, "A theory of falling growth and rising rents," Technical Report, National Bureau of Economic Research 2019.

_ , Ufuk Akcigit, and Peter Howitt, "What do we learn from Schumpeterian growth theory?," in "Handbook of economic growth," Vol. 2, Elsevier, 2014, pp. 515–563.

_ , _ , Antonin Bergeaud, Richard Blundell, and David Hémous, "Innovation and top income inequality," *The Review of Economic Studies*, 2018, *86* (1), 1–45.

_ , _ , Matthieu Lequien, and Stefanie Stantcheva, "Monetary Incentives, Tax Evasion and the Quest for Simplicity," 2019. mimeo College de France.

Akcigit, Ufuk and Stefanie Stantcheva, "Taxation and Innovation: What Do We Know?," Working Paper 27109, National Bureau of Economic Research May 2020.

_ and William R Kerr, "Growth through heterogeneous innovations," *Journal of Political Economy*, 2018, *126* (4), 1374–1443.

Alesina, Alberto, Michele Battisti, and Joseph Zeira, "Technology and labor regulations: theory and evidence," *Journal of Economic Growth*, 2018, *23* (1), 41–78.

Autor, David H, William R Kerr, and Adriana D Kugler, "Does employment protection reduce productivity? Evidence from US states," *The Economic Journal*, 2007, *117* (521), F189–F217.

Axtell, Robert L, "Zipf distribution of US firm sizes," *science*, 2001, *293* (5536), 1818–1820.

Banerjee, Abhijit and Esther Duflo, "Growth Theory through the lens of development," in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Amsterdam: Elsevier, 2005, chapter 7, pp. 474–544.

Barlevy, Gadi, "On the cyclicality of research and development," *American Economic Review*, 2007, *97* (4), 1131–1164.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta, "Cross-country differences in productivity: The role of allocation and selection," *The American Economic Review*, 2013, *103* (1), 305–334.

Bartelsman, Eric J, Pieter A Gautier, and Joris De Wind, "Employment protection, technology choice, and worker allocation," *International Economic Review*, 2016, *57* (3), 787–826.

Bassanini, Andrea, Luca Nunziata, and Danielle Venn, "Job protection legislation and productivity growth in OECD countries," *Economic policy*, 2009, *24* (58), 349–402.

Belenzon, Wes Cohen Ashish Arora Sharon and Honggi Lee, "Big Firms and the Direction of Technical Change," 2020. Mimeo Duke University.

Bena, Jan, Hernan Ortiz-Molina, and Elena Simintzi, "Shielding firm value: Employment protection and process innovation.," 2020. mimeo University of British Columbia.

Bento, Pedro and Diego Restuccia, "Misallocation, establishment size, and productivity," *American Economic Journal: Macroeconomics*, 2017, *9* (3), 267–303.

Bentolila, Samuel and Giuseppe Bertola, "Firing costs and labour demand: how bad is eurosclerosis?," *The Review of Economic Studies*, 1990, *57* (3), 381–402.

Bergeaud, Antonin, Yoann Potiron, and Juste Raimbault, "Classifying patents based on their semantic content," *PloS one*, 2017, *12* (4), e0176310.

Besley, Timothy and Robin Burgess, "Can Labor Regulation Hinder Economic Performance? Evidence from India," *Quarterly Journal of Economics*, 2000, *119* (1), 91–134.

Blundell, Richard, Rachel Griffith, and Frank Windmeijer, "Individual effects and dynamics in count data models," *Journal of econometrics*, 2002, *108* (1), 113–131.

Boedo, Hernan J Moscoso and Toshihiko Mukoyama, "Evaluating the effects of entry regulations and firing costs on international income differences," *Journal of Economic Growth*, 2012, *17* (2), 143–170.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-Experimental Shift-Share Research Designs," Working Paper 24997, National Bureau of Economic Research September 2018.

Braguinsky, Serguey, Lee G Branstetter, and Andre Regateiro, "The Incredible Shrinking Portuguese Firm," Working Paper 17265, National Bureau of Economic Research July 2011.

Ceci-Renaud, Nila and Paul-Antoine Chevalier, "L'impact des seuils de 10, 20 et 50 salariés sur la taille des entreprises Françaises," *Economie et Statistique*, 2011, *437*, 29–45.

Cette, Gilbert, Jimmy Lopez, and Jacques Mairesse, "Labour Market Regulations and Capital Intensity," Working Paper 22603, National Bureau of Economic Research September 2016.

_ , Lorraine Koehl, and Thomas Philippon, "Labor Shares in Some Advanced Economies," Working Paper 26136, National Bureau of Economic Research August 2019.

Chetty, Raj, John Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *Quarterly Journal of Economics*, 2011, *126* (2), 749–804.

Crépon, Bruno and Emmanuel Duguet, "Estimating the innovation function from patent numbers: GMM on count panel data," *Journal of Applied Econometrics*, 1997, *12* (3), 243–263.

Da-Rocha, José-María, Diego Restuccia, and Marina Mendes Tavares, "Firing costs, misallocation, and aggregate productivity," *Journal of Economic Dynamics and Control*, 2019, *98*, 60–81.

Davis, Steven J and John Haltiwanger, "Gross job creation, gross job destruction, and employment reallocation," *The Quarterly Journal of Economics*, 1992, *107* (3), 819–863.

Dechezlepretre, Antoine, David Hemous, Morten Olsen, and Carlo Zanella, "Automating labor: evidence from firm-level patent data," CEP Discussion Papers dp1679.pdf, Centre for Economic Performance, LSE February 2020.

_ , Elias Einiö, Ralf Martin, Kieu-Trang Nguyen, and John Van Reenen, "Do tax Incentives for Research Increase Firm Innovation? An RD Design for R&D," Working Paper 22405, National Bureau of Economic Research July 2016.

Gabler, Alain and Markus Poschke, "Experimentation by firms, distortions, and aggregate productivity," *Review of Economic Dynamics*, 2013, *16* (1), 26–38.

Garcia-Vega, Maria, Richard Kneller, and Joel Stiebale, "Labor Market reform and innovation: Evidence from Spain," Working Paper 201917, University of Nottingham 2019.

Garicano, Luis, Claire Lelarge, and John Van Reenen, "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, November 2016, *106* (11), 3439–79.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik instruments: What, when, why, and how," *American Economic Review*, 2020, *110* (8), 2586–2624.

Gourio, Francois and Nicolas Roys, "Size-Dependent Regulations, Firm size distribution and Reallocation," *Quantitative Economics*, 2014, *5* (2), 377–416.

Griffith, Rachel and Gareth Macartney, "Employment protection legislation, multinational firms, and innovation," *Review of Economics and Statistics*, 2014, *96* (1), 135–150.

Grout, Paul A, "Investment and wages in the absence of binding contracts: A Nash bargaining approach," *Econometrica: Journal of the Econometric Society*, 1984, pp. 449–460.

Gust, Christopher and Jaime Marquez, "International comparisons of productivity growth: the role of information technology and regulatory practices," *Labour economics*, 2004, *11* (1), 33–58.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "Market Value and Patent Citations," *RAND Journal of Economics*, 2005, *36* (1), 16–38.

Hall, Robert E, "The relation between price and marginal cost in US industry," *Journal of political Economy*, 1988, *96* (5), 921–947.

Hopenhayn, Hugo A, "On the Measure of Distortions," Working Paper 20404, National Bureau of Economic Research August 2014.

Hsieh, C. and P. Klenow, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 2009, pp. 1403–1448.

Hsieh, Chang-Tai and Benjamin A Olken, "The missing "missing middle"," *Journal of Economic Perspectives*, 2014, *28* (3), 89–108.

Hummels, David, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, "The wage effects of offshoring: Evidence from Danish matched worker-firm data," *American Economic Review*, 2014, *104* (6), 1597–1629.

INSEE, "NAF REV. 2 et CPF REV. 2 : Guide d'utilisation," 2016.

Jones, Charles I, "Misallocation, Economic Growth, and Input-Output Economics," Working Paper 16742, National Bureau of Economic Research 2011.

Kaplow, Louis, "Optimal regulation with exemptions and corrective taxation," 2013. Harvard University mimeo.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring technological innovation over the long run," Technical Report w25266, National Bureau of Economic Research 2018.

Klette, Tor Jakob and Samuel Kortum, "Innovating firms and aggregate innovation," *Journal of political economy*, 2004, *112* (5), 986–1018.

Kleven, Henrik and Mazhar Waseem, "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *Quarterly Journal of Economics*, 2013, *128* (2), 669–723.

Kramarz, Francis and Marie-Laure Michaud, "The shape of hiring and separation costs in France," *Labour Economics*, 2010, *17* (1), 27–37.

Lequien, Matthieu, Loriane Py, and Paul Trichelair, "Linking patents to firms: insights with French firms," 2017. mimeo.

Loecker, Jan De, Jan Eeckhout, and Gabriel Unger, "The rise of market power and the macroeconomic implications," *The Quarterly Journal of Economics*, 2020, *135* (2), 561–644.

Lucas, Robert E, "On the size distribution of business firms," *The Bell Journal of Economics*, 1978, pp. 508–523.

Manera, Andrea and Martina Uccioli, "Employment Protection and the Direction of Technology Adoption," 2020. mimeo MIT.

Mann, Katja and Lukas Püttmann, "Benign effects of automation: New evidence from patent texts," 2018. Available at SSRN 2959584.

Manso, Gustavo, Benjamin Balsmeier, and Lee Fleming, "Heterogeneous Innovation and the Antifragile Economy," 2019. mimeo, UC Berkeley.

Mayer, Thierry, Marc J. Melitz, and Gianmarco I. P. Ottaviano, "Market Size, Competition, and the Product Mix of Exporters," *American Economic Review*, 2014, *104* (2), 495–536.

_ , _ , and _ , "Product Mix and Firm Productivity Responses to Trade Competition," NBER Working Papers 22433, National Bureau of Economic Research, Inc July 2016.

Menezes-Filho, Naercio, David Ulph, and John Van Reenen, "The determination of R&D: empirical evidence on the role of unions," *European Economic Review*, 1998, *42* (3-5), 919–930.

Moins, Laurent, *Code Du Travail 2010*, Lamy, 2010.

Mukoyama, Toshihiko and Sophie Osotimehin, "Barriers to reallocation and economic growth: the effects of firing costs," *American Economic Journal: Macroeconomics*, 2019, *11* (4), 235–70.

Parente, Stephen L. and Edward C. Prescott, *Barriers to Riches*, Cambridge: MIT Press, 2000.

Porter, Michael E and Claas Van der Linde, "Toward a new conception of the environment-competitiveness relationship," *Journal of economic perspectives*, 1995, *9* (4), 97–118.

Poschke, Markus, "Employment protection, firm selection, and growth," *Journal of Monetary Economics*, 2009, *56* (8), 1074–1085.

Restuccia, Diego and Richard Rogerson, "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, October 2008, *11* (4), 707–720.

Saez, Emmanuel, "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2010, *2* (3), 180–212.

Saint-Paul, Gilles, "Employment protection, international specialization, and innovation," *European Economic Review*, 2002, *46* (2), 375–395.

Samaniego, Roberto M, "Employment protection and high-tech aversion," *Review of Economic Dynamics*, 2006, *9* (2), 224–241.

Schivardi, Fabiano and Tom Schmitz, "The IT revolution and southern Europe's two lost decades," *Journal of the European Economic Association*, 2020, *18* (5), 2441–2486.

Schumpeter, Josef, *Business Cycles*, New York: McGraw Hill, 1939.

Shleifer, Andrei, "Implementation cycles," *Journal of Political Economy*, 1986, *94* (6), 1163–1190.

Smagghue, Gabriel, "Heterogeneous Policy Distortions and the Labor Share," 2020. mimeo Banque de France.

Srebrovic, Rob, "Expanding your patent set with ML and BigQuery," 2019. Google Cloud Data Analytics https://cloud.google.com/blog/products/data-analytics/expanding-your-patent-set-with-ml-and-bigquery.

Webb, Michael, "The impact of artificial intelligence on the labor market," 2019. Available at SSRN 3482150.

ONLINE APPENDICES
NOT INTENDED FOR PUBLICATION

A More Details of some Size-Related Regulations in France

The size-related regulations are defined in four groups of laws. The Code du Travail (labor laws), Code du Commerce (commercial law), Code de la Sécurité Social (social security) and in the Code General des Impots (fiscal law). The main bite of the labor (and some accounting) regulations comes when the firm reaches 50 employees. But there are also some other size-related thresholds at other levels. The main other ones comes at 10-11 employees. For this reason we generally trim the analysis below 10 employees to mitigate any bias induced in estimation from these other thresholds. For more details on French regulation see inter alia Abowd and Kramarz (2003) and Kramarz and Michaud (2010), or, more administratively and exhaustively, Moins (2010).

A.1 Main Labor Regulations

The unified and official way of counting employees has been defined since 2004[34] in the Code du Travail,[35] articles L.1111-2 and 3. Exceptions to the 2004 definition are noted in parentheses in our detailed descriptions of all the regulations below. Employment is taken over a reference period which from 2004 was the calendar year (January 1st to December 31st). There are precise rules over how to fractionally count part-year workers, part-time workers, trainees, workers on sick leave, etc. (Moins, 2010). For example, say a firm employs 10 full-time workers every day but in the middle of the year all 10 workers quit and are immediately replaced by a different 10 workers. Although in the year as a whole 20 workers have been employed by the firm the standard regulations would mean the firm was counted as 10 employee firm. In this case, this would be identical to the concept used in our main FICUS data. Garicano et

[34]Before that date, the concept of firm size was different across labor regulations.
[35]The text is available at the legifrance website

al. (2016) extensively document that the discontinuity in the firm size distribution at 50 can be seen across a variety of firm datasets with different definitions of employment (e.g. DADS). There is of course more measurement error in some datasets than others due to differences in how the employment concept matches the regulatory definitions.

Recall that the employment measure in the FICUS data is average headcount number of employees taken on the last day of each quarter in the fiscal year (usually but not always ending on December 31st). All of these regulations strictly apply to the firm level, which is where we have the FICUS data. Some case law has built up, however, which means that a few of them are also applied to the group level.

From 200 employees:

- Obligation to appoint nurses (Code du Travail, article R.4623-51)
- Provision of a place to meet for union representatives (Code du Travail, article R.2142-8)

From 50 employees:

- Monthly reporting of the detail of all labor contracts to the administration (Code du Travail, article D.1221-28)
- Obligation to establish a staff committee ("comité d'entreprise") with business meeting at least every two months and with minimum budget = 0.3% of total payroll (Code du Travail, article L.2322-1-28, threshold exceeded for 12 months during the last three years)
- Obligation to establish a committee on health, safety and working conditions (CHSC) (Code du Travail, article L.4611-1, threshold exceeded for 12 months during the last three years)
- Appointing a shop steward if demanded by workers (Code du Travail, article L.2143-3, threshold exceeded for 12 consecutive months during the last three years)
- Obligation to establish a profit sharing scheme (Code du Travail, article L.3322-2, threshold exceeded for six months during the accounting year within one year after the year end to reach an agreement)

- Obligation to do a formal "Professional assessment" for each worker older than 45 (Code du Travail, article L.6321-1)
- Higher duties in case of an accident occurring in the workplace (Code de la Sécurité sociale and Code du Travail, article L.1226-10)
- Obligation to use a complex redundancy plan with oversight, approval and monitoring from Ministry of Labor in case of a collective redundancy for 9 or more employees (Code du Travail, articles L.1235-10 to L.1235-12; threshold based on total employment at the date of the redundancy)

From 25 employees:

- Duty to supply a refectory if requested by at least 25 employees (Code du Travail, article L.4228-22)
- Electoral colleges for electing representatives. Increased number of delegates from 25 employees (Code du Travail, article L.2314-9, L.2324-11)

From 20 employees:

- Formal house rules (Code du Travail, articles L.1311-2)
- Contribution to the National Fund for Housing Assistance;
- Increase in the contribution rate for continuing vocational training of 1.05% to 1.60% (Code du Travail, articles L.6331-2 and L.6331-9)
- Compensatory rest of 50% for mandatory overtime beyond 41 hours per week

From 11 employees:

- Obligation to conduct the election of staff representatives (threshold exceeded for 12 consecutive months over the last three years) (Code du Travail, articles L.2312-1)

From 10 employees:

- Monthly payment of social security contributions, instead of a quarterly payment (according to the actual last day of previous quarter);
- Obligation for payment of transport subsidies (Article R.2531-7 and 8 of the General Code local authorities, Code general des collectivités territoriales);
- Increase the contribution rate for continuing vocational training of 0.55% to 1.05% (threshold exceeded on average 12 months).

Note that, in additions to these regulations, some of the payroll taxes are related to the number of employees in the firm.

A.2 Accounting rules

The additional requirements depending on the number of employees of entreprises, but also limits on turnover and total assets are as follows (commercial laws, Code du Commerce, articles L.223-35 and fiscal regulations, Code général des Impôts, article 208-III-3):

From 50 employees:

- Loss of the possibility of a simplified presentation of Schedule 2 to the accounts (also if the balance sheet total exceeds 2 million or if the CA exceeds 4 million);
- Requirement for LLCs, the CNS, limited partnerships and legal persons of private law to designate an auditor (also if the balance sheet total exceeds 1.55 million euros or if the CA is more than 3.1 million euros, applicable rules of the current year).

From 10 employees:

- Loss of the possibility of a simplified balance sheet and income statement (also if the CA exceeds 534 000 euro or if the balance sheet total exceeds 267 000 euro, applicable rule in case of exceeding the threshold for two consecutive years).

B Data Appendix

B.1 Patent data

Our first database is PATSTAT Spring 2016's version which contains detailed information about patent applications from every patent office in the world. Among the very rich set of information available, one can retrieve the date of application, the technological class, the name of the patent holder (the *assignee*, the entity which owns the intellectual property rights) and the complete list of forward and backward citations.

We use a crosswalk built by Lequien et al. (2017) that associates each patent whose assignee is located in France with the official identifying number (or *SIREN*), which enables us to use most administrative firm level datasets. This matching use supervised learning based on a training sample of manually matched patents from the French patent office (INPI). It has the advantage over other matching protocols as it is specific to French firms to exploits additional information such as the location of innovative establishments (see Lequien et al., 2017 or Aghion et al., 2018a for more details).[36]

Because we stop our patent analysis in 2007, we are not affected by the truncation bias toward the end of the sample (see Hall et al., 2005) and we consider that our patent information are complete. In order to be as close to the time of the innovation as possible, we follow the literature and consider the filing year and not the granting year in our study. We use citations through to the last year (2016). When calculating a firm's quantile in the patent citation distribution, we do this based on a technology class (32 codes) by cohort-year.

Finally, we consider every patent owned by a French firm, regardless of the patent office that granted the patent rights, but we restrict to *priority patents* which correspond to the earliest patents which relate to the same invention. Therefore, if a firm successively fills the same patent in different patent offices, only the first application of this family will be counted.

[36] If the firm shares a patent with another firm, then we only allocate a corresponding share of this patent to the firm.

B.2 Firm-level administrative data

Our second data source provides us with accounting data for French firms from the DGFiP-INSEE, this data source is called `FICUS`. The data are drawn from compulsory reporting of firms and income statements to fiscal authorities in France. Since every firm needs to report every year to the tax authorities, the coverage of the data is all French firms from 1994 to 2007 with no limiting threshold in terms of firm size or sales. This dataset provides us with information on the turnover, employment, value-added, the four-digit NACE sector the firm belongs to. This corresponds to around 35 million observations.

The manufacturing sector is defined as category C of the first level of the NAF (*Nomenclature d'Activités Françise*), the first two digits of which are common to both NACE (Statistical Classification of Economic Activities in the European Community) and ISIC (International Standard Industrial Classification of All Economic Activities). INSEE provides each firm with a detailed principal activity code (APE) with a top-down approach: it identifies the 1-digit section with the largest value added. Among this section, it identifies the 2-digit division with the largest value-added share, and so on until the most detailed 5-digit APE code (INSEE, 2016). It is therefore possible that another 5-digit code shows a larger value-added share than the APE identified, but one can be sure that the manufacturing firms identified produce a larger value-added in the manufacturing section than in any other 1-digit section, which is precisely what we rely on to select the sample of most of our regressions. The 2-digit NAF sector, which we rely intensively on for our fixed effects, then represents the most important activity among the main section of the firm. Employment each year is measured on average within the year and may therefore be a non-integer number.

B.3 Trade data

Customs data for French firms Detailed data on French exports by product and country of destination for each French firm are provided by French Customs. These are the same data as in Mayer et al. (2014) but extended to the whole 1994-2012 period. Every firm must report its exports by destination country and by very detailed product (at a level finer than HS6). However administrative simplifications for intra-EU trade have been implemented since the Single Market, so that when a firm annually exports inside the EU less than a given threshold, these intra-EU flows are not reported and therefore not in our dataset. The threshold stood at 250 000 francs in 1993, and has been periodically reevaluated (650 000 francs in 2001, 100 000 euros in 2002, 150 000 euros in 2006). Furthermore flows outside the EU both lower than 1 000 euros in value and 1 000 kg in weight are also excluded until 2009, but this exclusion was deleted in 2010.

Country-product bilateral trade flows CEPII's database `BACI`, based on the UN database `COMTRADE`, provides bilateral trade flows in value and quantity for each pair of countries from 1995 to 2015 at the HS6 product level, which covers more than 5,000 products. To convert HS products into ISIC industries we use a United Nations correspondence table (when 1 HS code corresponds to 2 ISIC codes, we split the HS flow in half into each ISIC code).

C Theoretical Appendix

C.1 Solving numerically the baseline model

We solve the model numerically. To do so, we need to discretize the problem. That is, we need to move from a model with a continuum of products of size 1 to a model with a finite number of products K and a finite number of firms N.

The final good aggregator is adjusted as follows:

$$\ln y = \int_0^1 \ln y_j dj \text{ becomes } \ln y = \frac{1}{K}\sum_{j=1}^{K} \ln y_j$$

Unite price of a given intermediate good j is unchanged, but the demand:

$$y_j = \frac{y}{p_j} \text{ becomes } y_j = \frac{y}{p_j K}$$

And as a result:

$$\pi_j = \left(1 - \frac{1}{\gamma}\right) y \text{ becomes } \pi_j = \left(1 - \frac{1}{\gamma}\right) \frac{y}{K}$$

Finally, firm i's employment L_i is still equal to $n/(\omega\gamma)$ where:

$$\omega = \frac{wK}{y}.$$

The firm's maximization problem is still:

$$n\pi(n) + \beta nz\left[(n+1)\pi(n+1) - n\pi(n)\right] + \beta nx\left[(n-1)\pi(n-1) - n\pi(n)\right] - \zeta z^{\eta} n \frac{y}{K}$$

where the R&D cost function in our finite line model

$$C(z, n) = \zeta n z^{\eta} y \text{ has become } C(z, n) = \frac{\zeta}{K} n z^{\eta} y$$

With these changes in mind, equation (4) still applies and we can numerically solve the model in steady state. We proceed as follows:

1. There is a finite number N of firms and K of product lines, with $K > N$
2. $\mu(n)$ denotes the number of firms producing in exactly n product lines and $z(i)$ denotes its innovation intensity per line (which is taken from equation (4) in the model).
3. All firms produce at least one product, as a result, we must have $\mu(n) = 0$ for all $n \geq K - N$. For all i larger than 1

We therefore have $K - N + 1$ unknowns: $\mu(n)$ for $1 \leq n < K - N$ ($K - N - 1$ unknowns), x and z_e. The corresponding $K - N + 1$ independent equations are given by:

- The law of motion for μ:

$$\mu(n) = \frac{(n-1)\mu(n-1)z(n-1) + \mu(n+1)(n+1)x}{n(x + z(n))},$$

  for all $n \geq 2$ and $n < K - N$, recalling that $\mu(K - N) = 0$

- The definition of μ:

$$\sum_{n=1}^{K-N-1} \mu(n) = N$$

- The definition of x

$$x = z_e + \sum_{n=1}^{K-N-1} z(n)n\mu(n)/K$$

- The steady-state equation for the number of firms in the economy

$$\mu(1)x = z_e K$$

C.2 Theoretical extensions

C.2.1 Radical vs. Incremental innovation

Innovation equation: we solve for $u(n)$ and $z(n)$ by taking the first order condition from equation (9), where z is the output-adjusted effort invested in incremental R&D and u is the output-adjusted effort invested in radical R&D. This yields the following two equations:

$$u(n) = \left(\frac{\beta}{\alpha\eta}\left[(n+k)\pi(n+k) - (n+1)\pi(n+1)\right]\right)^{\frac{1}{\eta-1}}$$

and

$$z(n) = \left(\frac{\beta}{\zeta\eta}\left[(n+1)\pi(n+1) - n\pi(n)\right]\right)^{\frac{1}{\eta-1}} - \left(\frac{\beta}{\alpha\eta}\left[(n+k)\pi(n+k) - (n+1)\pi(n+1)\right]\right)^{\frac{1}{\eta-1}}$$

With these two expressions, we can solve for the equilibrium size distribution and for the share of radical innovations over incremental innovations for each firm size. See Figures 7 and 8.

The equilibrium size distribution is depicted in Figure C1, first on a linear scale and then on a logarithmic scale. The size distribution is qualitatively similar to the baseline case.

A special case when $\eta = 2$: We solve formally for u and z in equation (9) in the simple case where we take the overall cost of R&D to be quadratic and equal to $\zeta(u+z)^2 n/2 + \alpha u^2 n/2$. Thanks to the quadratic cost assumption, the first-order conditions can be conveniently summarized by the linear system:

$$\begin{pmatrix} \zeta & \zeta \\ \zeta & \alpha+\zeta \end{pmatrix} \begin{pmatrix} z \\ w \end{pmatrix} = \beta \begin{pmatrix} (n+1)\pi(n+1) - n\pi(n) \\ (n+k)\pi(n+k) - n\pi(n) \end{pmatrix}$$

As long as α and ζ are not equal to 0, this linear system solves into:

$$\begin{pmatrix} z \\ w \end{pmatrix} = \frac{\beta}{\zeta\alpha} \begin{pmatrix} \zeta+\alpha & -\zeta \\ -\zeta & \zeta \end{pmatrix} \begin{pmatrix} (n+1)\pi(n+1) - n\pi(n) \\ (n+k)\pi(n+k) - n\pi(n) \end{pmatrix}$$

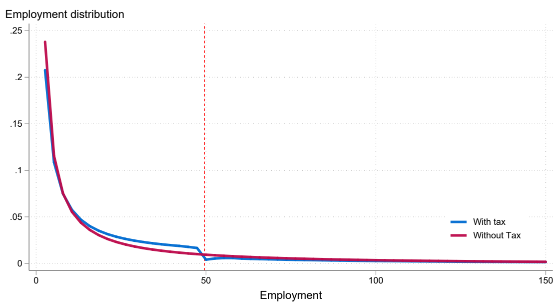
The solutions are presented in Table C1, where we have defined $\pi \equiv \frac{\gamma-1}{\gamma}\beta$

	$z(n)$	$u(n)$
$n < \bar{n} - k$	$\frac{\pi}{\alpha\zeta}\left(\alpha - \zeta(k-1)\right)$	$\frac{\pi}{\alpha}(k-1)$
$\bar{n} - k \leq n < \bar{n} - 1$	$\frac{\pi}{\alpha\zeta}\left(\alpha - \zeta(k-1) + \zeta\tau(n+k)\right)$	$\frac{\pi}{\alpha}\left((k-1) - \tau(k+n)\right)$
$n = \bar{n} - 1$	$\frac{\pi}{\alpha\zeta}\left(\alpha - \zeta(k-1) - \zeta\tau(\bar{n}-k) - \alpha\tau\bar{n}\right)$	$\frac{\pi}{\alpha}\left((k-1) + \tau(\bar{n}-k)\right)$
$n \geq \bar{n}$	$\frac{\pi}{\alpha\zeta}(1-\tau)\left(\alpha - \zeta(k-1)\right)$	$\frac{\pi}{\alpha}(1-\tau)(k-1)$
	$u(n) + z(n)$	$\frac{u(n)}{z(n)+u(n)}$
$n < \bar{n} - k$	$\frac{\pi}{\zeta}$	$\frac{\zeta}{\alpha}(k-1)$
$\bar{n} - k \leq n < \bar{n} - 1$	$\frac{\pi}{\zeta}$	$\frac{\zeta}{\alpha}(k-1)\left(1 - \frac{\tau(k+n)}{k-1}\right)$
$n = \bar{n} - 1$	$\frac{\pi}{\zeta}\left(1 - \tau\bar{n}\right)$	$\frac{\zeta}{\alpha}(k-1)\left(1 - \frac{\tau(\bar{n}-k)}{k-1}\right)\frac{1}{1-\tau\bar{n}}$
$n \geq \bar{n}$	$\frac{\pi}{\zeta}(1-\tau)$	$\frac{\zeta}{\alpha}(k-1)$

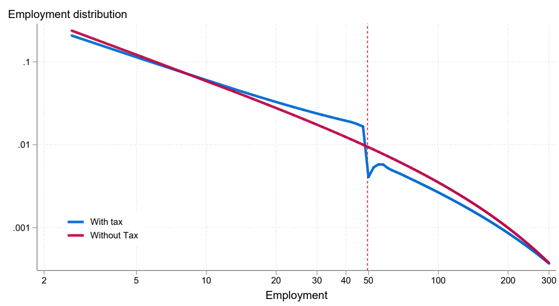
Table C1: Solution in the Extended Model with two types of innovation (radical and incremental)

Figure C1: Firm size distribution with two types of innovation

(a) Linear scale



(b) Log scale



Notes: These figures plot the value of $\mu(n)$ as a function of employment $L = n/(\gamma\omega)$. Left-hand side panel uses a linear scale and right-hand side panel a log-log scale. Extension with two types of innovation with $k = 4$ (see Section 5.1)

C.2.2 Infinitely Lived Owners

To solve numerically for the infinite-lived firm case, we proceed as follow. We start from equation (10) and derive the first order condition:

$$z(n) = \left(\frac{W(n+1) - W(n)}{\zeta\eta} \right)^{\frac{1}{\eta-1}} \text{ if } W(n+1) > W(n) \text{ and } 0 \text{ otherwise.}$$

This implies:

$$W(n-1) = \frac{1}{x}\left(W(n)(\frac{\rho}{n}+x) - \pi(n)\right) + \zeta z(n)^{\eta} - z(n+1)\left(W(n+1) - W(n)\right) \tag{C1}$$

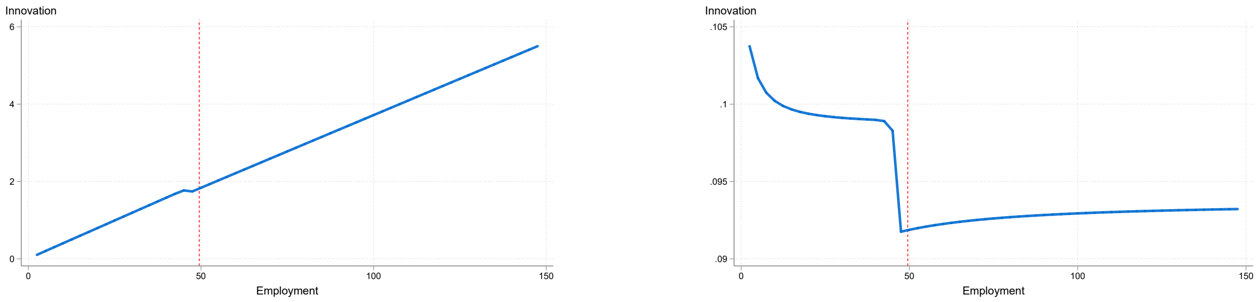
To solve recursively, we use a free entry construction which states that the monetary cost of entry is equal to $W(1)$ (see Aghion et al., 2014). We then solve backward:

- We guess a value for $W(K)$, where K is the number of product (see Appendix C.1). We know that $z(K) = 0$ are there is no incentive to invest in R&D given that a firm of size K cannot add more lines.
- We then use equation (C1) to find $W(K-1)$.
- We then solve for $z(K-1)$ and so on until we find $W(1)$
- We compare $W(1)$ to the entry cost. If the difference is larger in absolute value than some tolerance threshold, we start again with another guess for $W(K)$.

To simplify, we take x, the rate of creative destruction as exogenous but endogenizing it can be done in the same way as in the baseline case. Note also that we do not make any attempt at calibrating the entry cost. Hence, our results are mostly qualitative and are found in Figure C2 which depicts the relationship between size and innovation. Because the result is much smoother than in the baseline, we report both the value of total innovation ($nz(n)$) against size and innovation per line (z).

Figure C2: Innovation and size for infinitely-lived firms

(a) Total innovation (b) Innovation per line



Notes: These figures plot the value of $nz(n)$ (left-hand side) and $z(n)$ (right-hand side) as a function of employment $L = n/(\gamma\omega)$. In the extension of infinitely-lived firms.

C.2.3 R&D as scientific labor

This section solves the model outlaid in Section 5.4. In this extension, R&D is performed by scientists, hence the workforce is now split between production and innovation workers. For each firm i, employment L_i is therefore given by:

$$L_i = \frac{n_i}{\omega\gamma} + \zeta n_i z_i^\eta. \tag{C2}$$

Aggregating over all firms, we get:

$$\mathcal{L} = \int_i L_i di = \frac{1}{\omega\gamma} + \zeta \int_i n_i z_i^\eta di = \frac{1}{\omega\gamma} + \zeta \sum_{n>0} \mu(n) n z^\eta(n)$$

Since $\mathcal{L}$ is fixed and exogenous, and since the right hand side terms of the above equation varies with the tax τ, then ω also varies with τ. More precisely, the equilibrium wage ω decreases with τ, since regulation costs decreases aggregate innovation (the second term of the right-hand side of the equation).

Given that employment is now a function of both the number of products n and the intensity of innovation z, we denote it by $L(n,z)$. The cutoff threshold $\bar{l} = 50$ is now defined by the set of points in the space (n,z) such that:

$$z = \frac{1}{\zeta n}\left(\bar{l} - \frac{n}{\gamma\omega}\right) \tag{C3}$$

Figures C3 shows the equilibrium relationship between the number of products, employment and innovation intensity (which indirectly relates to the number of R&D workers). It is no longer possible to use the number of products as a measure of the size of the firms and we need to define profit per unit of final output is now equal to:

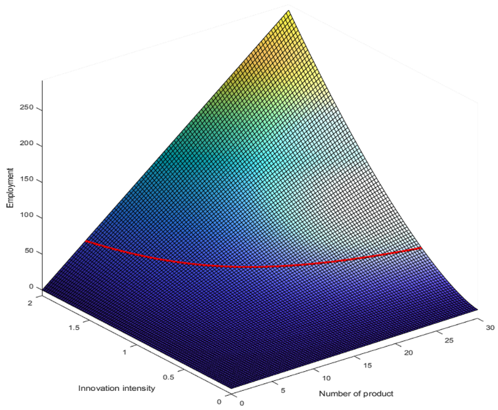
$$\pi(n,z) = \frac{\gamma - 1}{\gamma}\left(1 - \mathbb{1}\left[L(n,z) \geq \bar{l}\right]\tau\right).$$

Hence, the firm's maximization problem remains the same as before but with the two state variables n and z, that is:

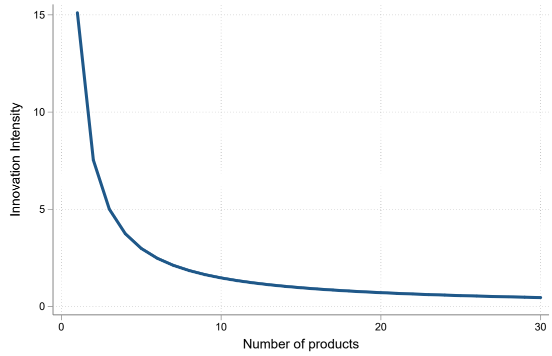
$$\max_{n\geq 0, z\geq 0}\left\{n\pi(n,z)y - \zeta n z^{\eta} y + \frac{1}{1+r}\mathbb{E}\left[n'\pi(n',z')y'\right]\right\}.$$

Figure C3: Localization of employment threshold $\bar{l}$

(a) 3D plot



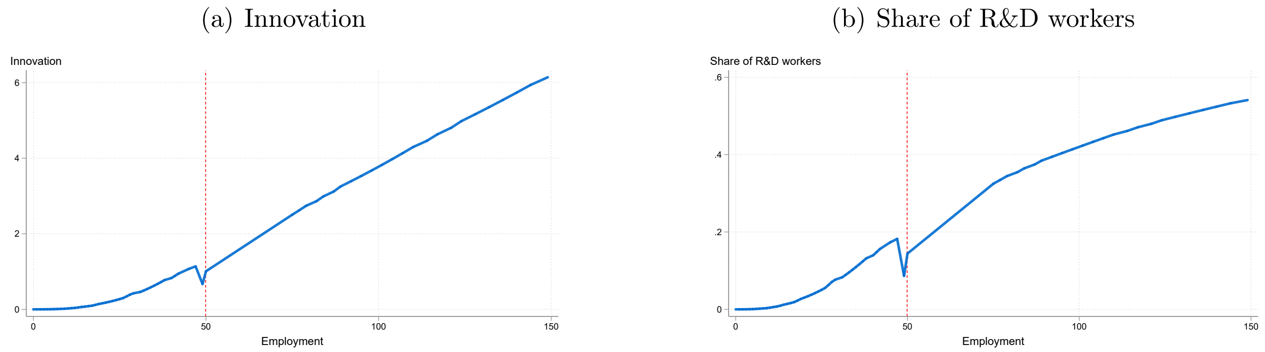
(b) 2D projection



Notes: These Figures plot the relationship between employment L, innovation intensity z and number of products n. The left-hand side panel shows the 3D plot corresponding to the surface defined by equation (C2), where the z-axis corresponds to L. The curve in red corresponds to the intersection of the surface (n, z, L) with the surface $L = \bar{l}$. The right-hand side panel presents the set of pairs (z, n) which corresponds to an employment level of $\bar{l}$ according to equation (C3).

Solving this maximization problem for every value of n gives a function $Z(n) = nz(n)$ which we plot in Figure C4 against employment $L(n)$. We see that the innovation-employment cross section relationship is qualitatively unchanged. In Figure C4, we also plot the corresponding relationship between firm's employment and its share of R&D workers.

Figure C4: Innovation-Employment cross-section with scientists $\bar{l}$



Notes: This is the total amount of innovation ($Z(n)$, left-hand side panel) and share of R&D workers in total employment ($\zeta n z^{\eta}(n)/L(n)$), right-hand side panel) by firms of different sizes (employment, $L = n/(\omega\gamma) + \zeta n z^{\eta}$) according to our theoretical model extension presented in Section 5.4. We use arbitrary parameter values for illustrative purposes.

D Additional Empirical Results

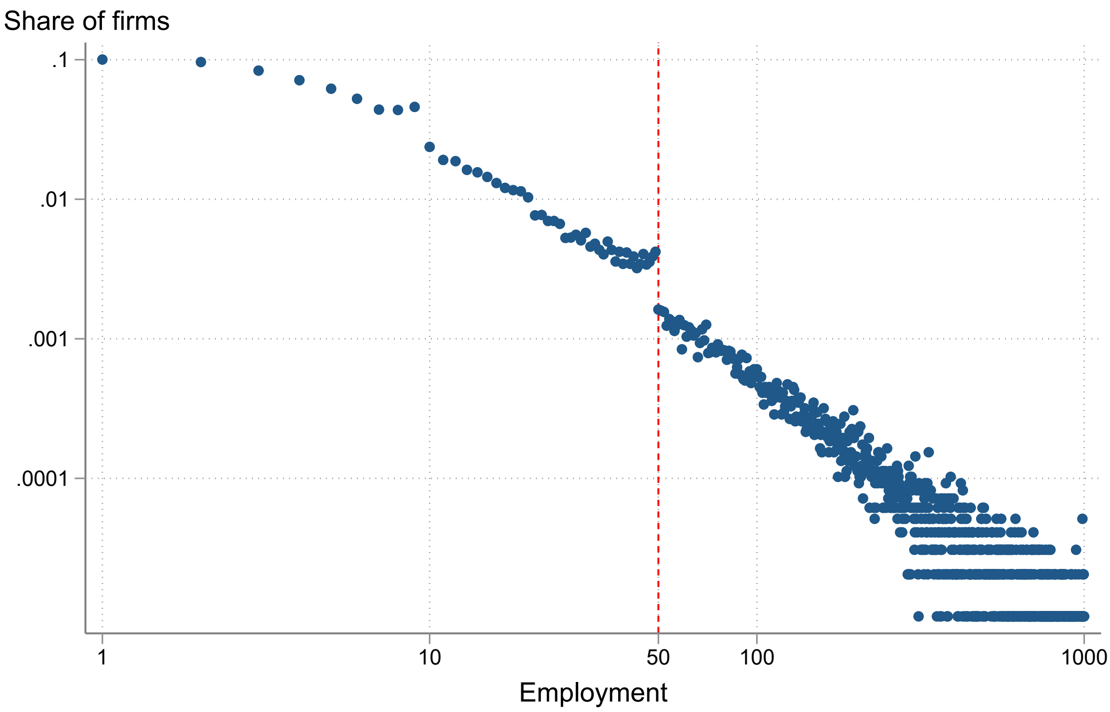
D.1 Size distribution of French firms

Figure D1 reports the size distribution of firms from FICUS in a log-log scale. In order to replicate results from Garicano et al. (2016), we use the year 2000, although choosing another year would results in a similar relationship. We also restrict to the manufacturing sector to be consistent with our empirical analysis. The relationship is consistent with the well-know power law documented namely by Axtell (2001), but with two discontinuities: one at 50 employees and the other one at 10 employees, corresponding to size dependent regulation thresholds (see Appendix A).

D.2 Robustness of the cross sectional innovation-size relationship

As noted in the main text the relationship between firm innovation and size are robust to a wide variety of alternative definitions. The baseline method in Figure 5 defines as innovative firm as one who has produced at least open patent over the sample period. In Figure D2 we consider using a narrower window around the year employment is measured. Panel A uses between patents in t, exactly the same year as employment as measured. Panel B uses patents filed two years before and two years after the employment measure (a five year window) and Panel C between four years before and after (a nine year window). Panel D measures innovation as log(1+patents count) in the same year as employment. Although the measures are somewhat

Figure D1: Distribution of size for the manufacturing sector



Notes: The data relate to the year 2000 for the manufacturing sector.

noisier than using the whole period (which smooths things out), the same basic pattern of an innovation valley and a fall in the gradient after the regulatory threshold are apparent.

Figure D3 repeats these four definitions for the Figures comparing incremental and radical patents as measured by future citations (analogous to Figure 9).

D.3 Robustness of the dynamic effects of the market size shock on innovation

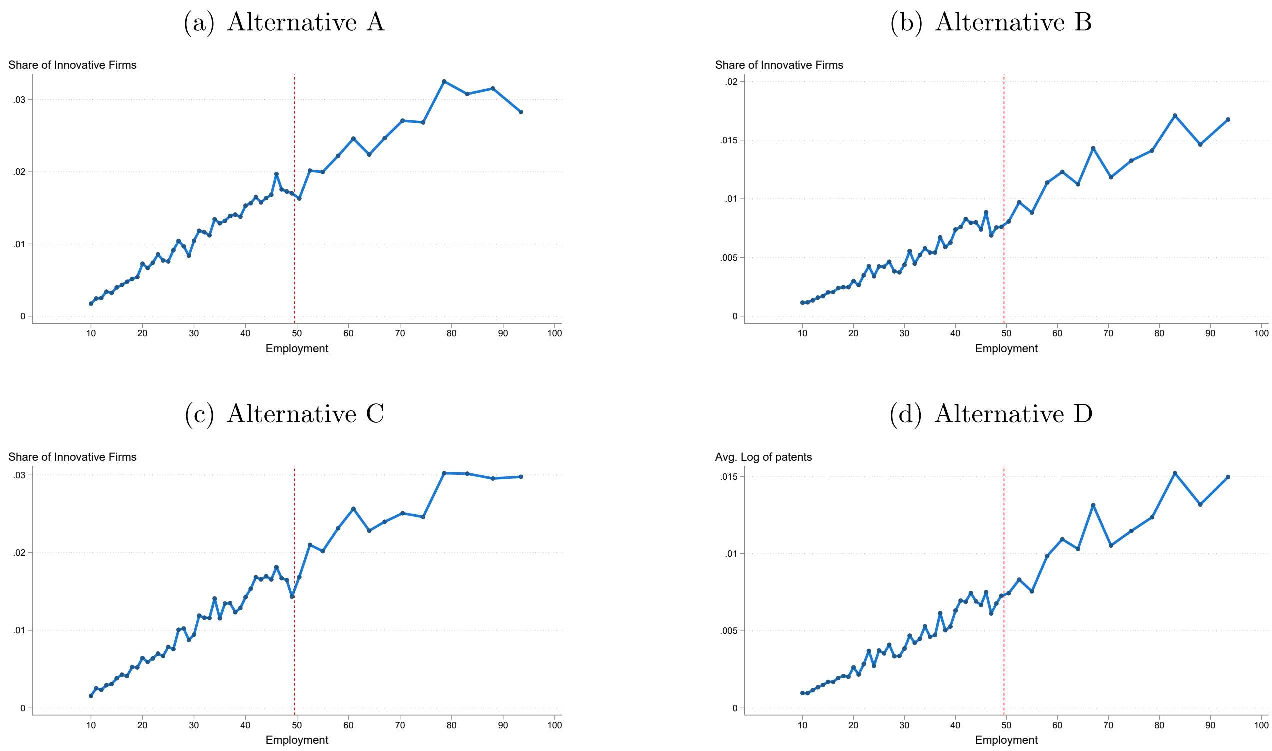
In the main text we noted the robustness of the decline in the impact of demand shocks to the left of the threshold and reported some of our tests. Here, we detail some more of these.

First, it is possible that the changing relationship between innovation and the market size shock around the threshold is driven by some kind of complex non-linearities in the innovation-employment relationship, and our quadratic controls are insufficient. To investigate this issue, we allow interactions between the demand shock and different size bins of firms in Table D1. Of all the 14 different size bins, only the interaction of the shock with the size bin just below the threshold (45-49 employees) is significantly different from zero and large in absolute magnitude.

Second, our results are robust to the particular way in which we define the upper and lower size cutoffs for our sample. Appendix Table D2 reproduces the baseline specification in column (1). Column (2) uses employment at t-2 instead of the initial year to define the sample, column (3) relaxes the upper threshold to include firms of up to 500 employees (instead of 100 employees in the baseline) and column (4) includes all firms below 100 employees (instead of dropping the firms with between zero and 9 workers). Column (5) restricts the sample to firms exporting in 1994 (instead of the restriction that a firm has to export in at least one year over the period 1994-2007). Column (6) includes all the non-exporting firms. The last three columns use three different definitions of the dependent variable instead of our basic measure $\tilde{\Delta} Y$: the log-difference in column (7), the difference in the Inverse Hyperbolic Sine in column (8) and the change in patents normalized on pre-sample patents in column (9). Our results are robust to all these tests.

Finally, one might be concerned that the quantiles of the citation distribution reported in Table 5 are arbitrary. Figure D4 reports the coefficient and confidence intervals on the key interaction term in our preferred specification for every quantile from the top 10th to the bottom 70th percentile in 5% intervals. As discussed in the text, it is clear that the negative effect of the regulation is only apparent for the less cited patents. There is no significant effect in a quantitative or statistical sense for patents in the top quartile of the citations distribution. The negative effect is driven by those in the bottom two-thirds of the citation distribution (with a monotonic decline of the effect for those between the 25th and 35th percentiles.

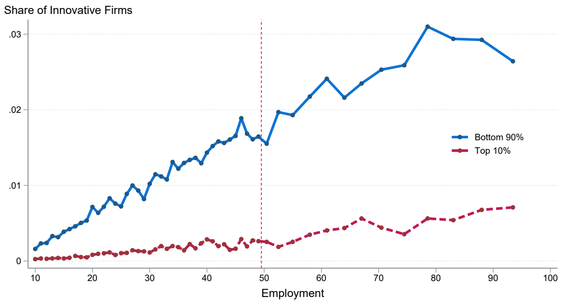
Figure D2: Innovative firms at each employment level - robustness



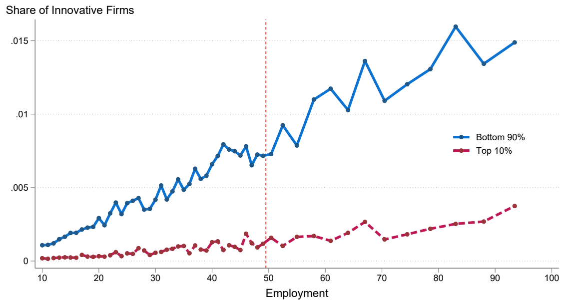
Notes: These Figures replicate Figure 5 using different definitions of the what counts as an innovative firm, based on the timing of patents. Alternatives A, B, C and D define an innovative firm as a firm having filed a priority patent application between $t-2$ and $t+2$ (A), at t (B), between $t-4$ and t (C). Alternative D uses the logarithm of 1 plus the number of patent application at t.

Figure D3: Innovative firms at each employment level and quality of innovation- robustness

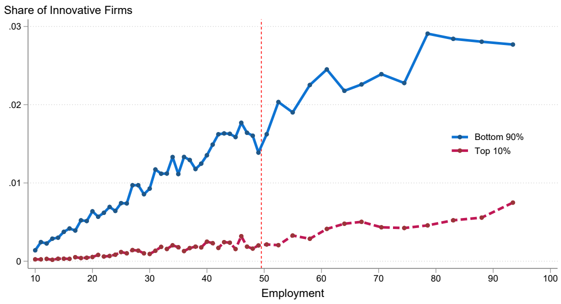
(a) Alternative A



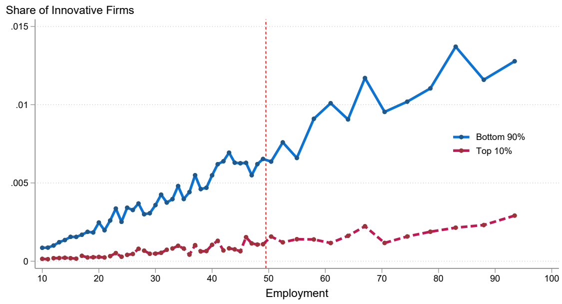
(b) Alternative B



(c) Alternative C

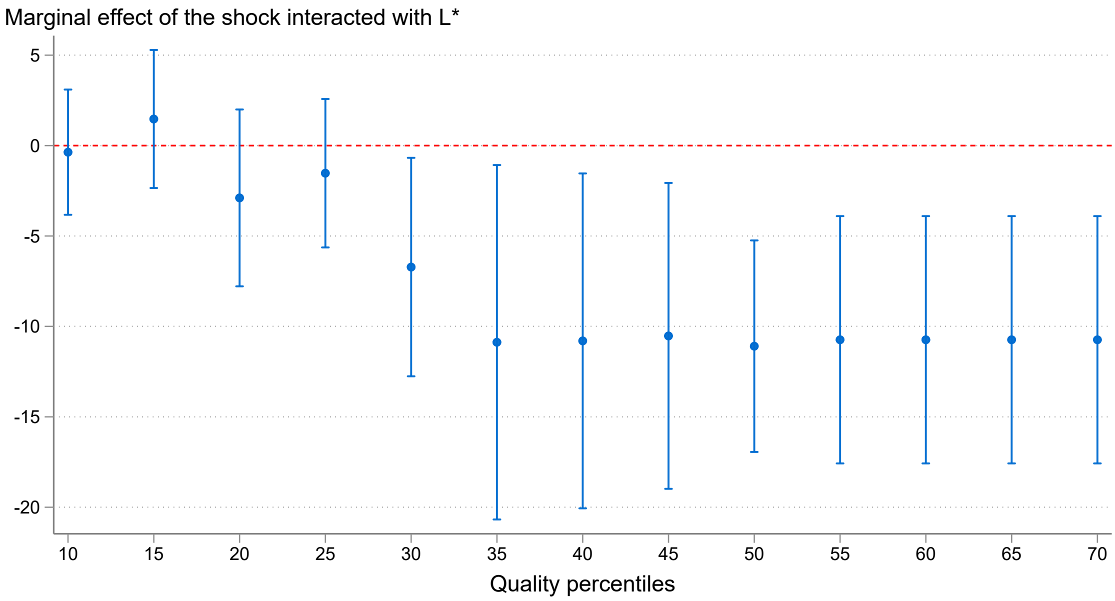


(d) Alternative D



Notes: These Figures replicate 9 using different definitions of the what counts as an innovative firm, based on the timing of patents. Alternatives A, B, C and D define an innovative firm as a firm having filed a priority patent application between $t-2$ and $t+2$ (A), at t (B), between $t-4$ and t (C). Alternative D uses the logarithm of 1 plus the number of patent application at t. The solid line considers the bottom 90% most cited patent and the dashed line the top 10% most cited.

Figure D4: Response to the Demand shock of patents of different quality



Notes: 95% confidence intervals around the estimated coefficient δ in equation (6). Each line corresponds to a separate estimation, where the dependent variable has been redefined by restricting to patents among the x% more cited in the year, with x equal to 10, 15 etc... up to 70. Note that the 65^{th} percentile threshold correspond to 0-citation patent and we include all patents for quality percentiles above 65. The estimated model is the same as in column 5 of Table 2.

Table D1: Placebo tests

	(1) 10-14	**(2)** 15-19	**(3)** 20-24	**(4)** 25-29	**(5)** 30-34	**(6)** 35-39	**(7)** 40-44	**(8)** **45-49**	**(9)** 50-54	**(10)** 55-59	**(11)** 60-64	**(12)** ≥ 65
$Shock_{t-2} \times L^{\star}_{t-2}$	-0.067	-1.077	0.475	-9.266*	2.385	6.237	4.584	**-10.713*** **	-6.638	5.955	5.106	0.417
	(1.761)	(2.293)	(5.751)	(5.210)	(5.861)	(5.605)	(6.266)	**(3.411)**	(5.921)	(7.473)	(7.389)	(6.189)
$L^{\star}_{t-2}$	0.062	-0.014	-0.038	0.130	-0.022	-0.099	0.117	**0.022**	-0.177	-0.126	-0.223	0.065
	(0.060)	(0.064)	(0.075)	(0.085)	(0.129)	(0.117)	(0.151)	**(0.098)**	(0.239)	(0.214)	(0.315)	(0.056)
$Shock_{t-2}$	-9.878*	-9.548**	-10.008**	-8.899**	-10.009**	-9.853**	-9.710**	**-10.818** **	-10.204**	-9.581**	-9.654**	-9.586*
	(4.917)	(4.522)	(4.468)	(4.221)	(4.350)	(4.378)	(4.471)	**(4.383)**	(4.320)	(4.329)	(4.356)	(5.541)
$log(L)_{t-2}$	-0.034	-0.045	-0.046	-0.043	-0.043	-0.041	-0.048	**-0.045**	-0.039	-0.041	-0.038	-0.063*
	(0.041)	(0.036)	(0.038)	(0.035)	(0.036)	(0.034)	(0.037)	**(0.035)**	(0.036)	(0.036)	(0.038)	(0.032)
$Shock_{t-2} \times log(L)_{t-2}$	3.609**	3.544**	3.633**	3.556**	3.589**	3.454**	3.454**	**4.104*** **	3.758**	3.472**	3.509**	3.505*
	(1.510)	(1.407)	(1.384)	(1.360)	(1.378)	(1.446)	(1.469)	**(1.424)**	(1.377)	(1.391)	(1.398)	(1.838)
Fixed Effects												
Sector×Year	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Number Obs.	191,010	191,010	191,010	191,010	191,010	191,010	191,010	**191,010**	191,010	191,010	191,010	191,010

Notes: These are based on the specification of column 5 of Table 2. The dependent variable is the Davis and Haltiwanger (1992) growth rate in the number of priority patent applications between $t-1$ and t. In each column $L^{\star}$ has been redefined as a dummy variable set to one if employment at $t-2$ is at different levels. These levels are defined as 10-14 (column 1), 15-19 (column 2), 20-24 (column 3) etc... up to over 65 (the baseline model is therefore in column 8). Innovation is measured by the number of new priority applications. . All models include a 3-digit NACE sector and a year fixed effects. Estimation period: 2007-1997 . Standard errors are clustered at the 3-digit NACE sector level. ***, ** and * indicate p-value below 0.01, 0.05 and 0.1 respectively.

Table D2: Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Shock_{t-2} \times L^{\star}_{t-2}$	-10.738***	-17.134***	-7.746*	-14.070**	-14.595**	-10.767***	-0.790***	-0.079*	-0.084*
	(3.419)	(5.544)	(4.327)	(5.399)	(5.399)	(3.364)	(0.236)	(0.038)	(0.041)
$L^{\star}_{t-2}$	0.026	-0.008	-0.010	0.002	-0.021	0.042	0.003	0.000	0.001
	(0.098)	(0.107)	(0.096)	(0.104)	(0.131)	(0.087)	(0.006)	(0.001)	(0.001)
$Shock_{t-2}$	-10.213**	-13.088	-8.690*	-7.404*	-13.955**	-10.131**	-0.727**	-0.092	-0.041
	(4.281)	(7.708)	(4.990)	(3.660)	(5.720)	(4.241)	(0.313)	(0.054)	(0.032)
$log(L)_{t-2}$	-0.054	-0.045	-0.030	-0.068*	-0.053	-0.041	-0.003	-0.000	-0.000
	(0.037)	(0.038)	(0.028)	(0.035)	(0.038)	(0.025)	(0.002)	(0.000)	(0.000)
$Shock_{t-2} \times log(L)_{t-2}$	3.972***	4.574*	3.533**	2.936**	5.127**	3.958***	0.283**	0.035*	0.017
	(1.387)	(2.400)	(1.585)	(1.330)	(1.853)	(1.382)	(0.100)	(0.018)	(0.011)
Fixed Effects									
Sector×Year	✓	✓	✓	✓	✓	✓	✓	✓	
Number Obs.	189,727	157,106	225,296	253,709	132,838	280,481	189,727	189,727	189,727

Notes: These are based on the specification of column 5 of Table 2. The dependent variable is the Davis and Haltiwanger (1992) growth rate in the number of priority patent applications between $t-1$ and t. Each column considers a different sample. Column (1) replicates our baseline specification. Column 2 includes firms that have a workforce between 10 and 100 employees at $t-2$ (instead of the first year they appear in the sample). Column 3 (resp. 4) includes firms that have a workforce between 10 and 500 (resp. 0 and 100) employees at t_0. Columns 5 and 6 are based on the same sample as column 1 but column 5 restricts to firm that first exported in 1994 (i.e.: $t_0 = 1994$, the earliest year in our dataset) and column 6 extends to non-exporting firms. Columns 7-9 also consider the same sample as column 1 but change the type of growth rate of the dependent variable. Column 7 considers the first difference in $log(1+Y)$, column 8 uses an hyperbolic function $log(Y+\sqrt{1+Y^2})$, also in first difference and column 9 uses the first difference of Y/S_0, where S_0 is the yearly average number of priority patents filed by the firm before t_0 (the first year the firm appears in the database) . All models include a 3-digit NACE sector and a year fixed effects. Estimation period: 2007-1997. Standard errors are clustered at the 3-digit NACE sector level. ***, ** and * indicate p-value below 0.01, 0.05 and 0.1 respectively.

D.4 Calculating Aggregate Innovation Loss: Alternative parameter estimates

In this section we give more details regarding the calculation of the aggregate innovation losses and test robustness to our main exercise in Section 4.

Our theoretical model predicts a relationship between Z and employment $l = n/(\gamma\omega)$. Specifically, equation (4) shows that:

$$Z \propto l \text{ if } l < (\bar{n} - 1)/(\gamma\omega) \text{ and } Z \propto l(1-\tau)^{\frac{1}{\eta-1}} \text{ if } l \geq \bar{n}/(\gamma\omega)$$

To map this into our data, we need to make an assumption on how Z relates to the number of patents filed by a firm. Our baseline estimates assume that $Z \propto P$, where P is the (smoothed) number of patent applications filed by the firm. We can therefore directly estimate τ from the slopes.

In this section, we make some robustness tests around these estimates. We report these in Table D3. Column (1) reports the baseline value and corresponding total innovation and welfare loss compared to an economy with $\tau = 0$.[37] Column (2) does the same but include an intercept, assuming that $Z = aP + b$ for some parameters a and b. Column (3) does the same

[37] To compute these loss, we have kept all other parameters the same. The only other parameters directly affected by changed in the slope estimates is β/ζ. However, we know that this parameter plays little aggregate role.

as column (1) but include firms up to 150 employees.[38]. Column (4) is the same as (1) but measures P using patents filed during the next three years and column (5) only uses patents filed during year $t+1$. Finally, columns (6), (7) and (8) assume that the relationship between Z and P depends on the sector and year. We thus perform an estimation without binning the data and include a sector fixed effect (column 6), a sector and year fixed effect (column 7) and a sector-year fixed effect (column 8).

Table D3: Alternative estimation of τ

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Observations	Employment binned					Firm level		
τ	2.5%	4.9%	1.1%	1.9%	1.4%	3.0%	3.6%	3.2%
Total Innovation loss (%)	5.42	1.14	2.31	4.05	2.95	6.59	8.05	7.07
Welfare loss (% of C equivalent)	2.16	4.53	0.92	1.62	1.18	2.63	3.20	2.82

Notes: This Table presents alternative OLS estimates of parameter τ based on the innovation-employment relationship of equation (4). τ is computed as the ratio of two slope, respectively for firms between 10 and 45 employees and for firms between 50 and 100 (except column 3 which extend this to 150). Left-hand side variable is the number of patents computed as a five year average around t (columns 1, 2, 3, 6, 7 and 8), in years $t+1$, $t+2$ and $t+3$ (column 4) and in year $t+1$ (column 5). Observations are either binned at the employment level (one observation per level of employment) in columns 1 to 5 or at the firm level and pooled together for other columns. Column 6 includes a 3-digit sector fixed effect, column 7 includes adds a year fixed effect (on top of the sector) and column 8 includes a interacted sector-year fixed effect. Each estimation includes dummies for each employment level between 46 and 49.

[38] Extending further the upper bound in the number of employees in our estimation samples reduces the value of the slopes for firms above the threshold. This is because the innovation-employment relationship flattens at some point. We however prefer to exclude large firms as we believe the relationship between Z and P is likely to change when firms become too large.

Table D4: Sensitivity analysis for welfare

Robustness	**Loss in total welfare**
Baseline	2.16%
$\gamma = 1.2$	0.78%
$\gamma = 1.50$	6.86%
$\eta = 2$	2.76%
$\eta = 1.3$	8.80%
$\omega = 0.22$	5.35%
$\omega = 0.29$	5.45%
$\beta/\zeta = 1.40$	1.57%
$\beta/\zeta = 1.90$	2.85%
$\beta = 0.94$	1.39%
$\beta = 0.98$	4.4%
τ	
Percentile 75^{th} ($\tau = 0.043$)	3.87%
Percentile 25^{th} ($\tau = 0.007$)	0.06%

Notes: baseline uses parameter values: ($\eta = 1.5$, $\gamma = 1.3$, $\tau = 0.025$, $\beta/\zeta = 1.65$ and $\omega = 0.25$), see Table 3. In the robustness where γ, η, ω or β/ζ are changed, we keep τ as in the baseline. The last two lines report the 25^{th} and 75^{th} percentile for the loss of innovation in a sample computed from 100,000 independent draws of τ from two normal distribution. The corresponding value of τ and β/ζ are computed as an average for each percentile. Loss in welfare is given in consumption equivalent and does not include initial quality (see section 4.3).

D.5 Measuring different types of innovation

Our baseline approach simply uses patent counts. In the extensions of C.2.1, we take several approaches to examining the different types of innovation. In order to measure how "radical" a patent is, we use two alternative methods: citations and text-based measures of novelty. Then we also consider measures of how "labor saving" the patent is by looking at measures of automation and process innovation.

Citations The first method uses the now classical approach of considering future citations. For every patent in a technology class by year of application cell, we calculate all the citations to that patent by all granted patents that were filed in 2016 or earlier. Since the last year we use in our analysis sample is 2007, this gives us a minimum of 10 future years of citation information. We then calculate which quantile of the citations distribution a given patent lies in. A patent which was in the top decile of citations, for example, would be counted as radical for the purposes of column 1 of Table 5.

We validated the use of this measure by presenting employment growth regressions. We regressed the change in the firm's log(employment) on a distributed lag of patent counts with sector by time dummies. Table D5 shows a representative example where we use patents from $t-1$ to $t-3$. To deal with zeros we add one to the patents before taking logs. Column (1) counts only "radical" patents in the top 10% of the technology-class-year cohort citation distribution and column (2) has incremental patents in the bottom 90%. The coefficients of all patents are positive and individually and jointly significant, indicating that patenting is associated with faster firm growth as we would expect. And consistent with our priors, the coefficients on the radical patents are much larger than incremental patents. Summing the coefficients to show the long-run effects in the base of the column we see that the radical patents have about 2.5 (=0.1402/0.0565) times the impact on employment growth compared to incremental patents. The base of the columns shows that in the long-run a doubling of incremental patents increases employment growth by 5.7% compared to 15% for radical patents.

More ambitiously, we can use these estimates to perform a back of the envelope calculation to see how much lower the loss of growth would be if we took into account that the regulation only affects incremental innovation. For example, using the approach of Table D5 radical innovations (the top 10% of the citations distribution) have 2.5 times the effect of incremental patents, so

Table D5: Regression results for different levels of the novelty of innovation

	Top 10%	Bottom 90%
$\log(P_{f,t-1})$	0.0598***	0.0187***
	(0.01524)	(0.00536)
$\log(P_{f,t-2})$	0.0466***	0.0223***
	(0.01472)	(0.00556)
$\log(P_{f,t-3})$	0.0338*	0.0154*
	(0.01742)	(0.00762)
Sum of coefficients	0.1402***	0.0565***
	(0.0210)	(0.0110)
Obs	196,284	196,284
R^2	0.0081	0.0093

Notes: The dependent variable is the change in the firm's log(employment). The left hand side is ln(1+patent count) between $t-1$ and $t-3$. Column (1) restricting to the top 10% most highly cited patents in a technology-class year and column (2) has the other 90% . Both models include a 2-digit NACE sector interacted with year fixed effects. Standard errors are clustered at the firm level. ***, ** and * indicate p-value below 0.01, 0.05 and 0.1 respectively.

an innovation index should give a weight of 5/7 to radical innovation and 2/7 to incremental innovation (instead of implicitly equal weights using the patent count). If the overall fall in patenting is 5.42% as estimated in Table 3 and this comes entirely from incremental innovation, we need to scale down the growth effect by 18/23 reflecting the lower impact of incremental innovation. This would imply a fall of 4.2% in growth (compared to the unregulated economy) compared to 5.4% in our baseline estimates. So the extension to different types of innovation does reduce the magnitude of the loss, but not by an enormous amount. Different assumptions will obviously change these exact magnitudes, but are unlikely (in our view) to overturn our main findings.

Our second, alternative measure of radical innovation involves a text-based analysis of novelty which is more involved and draws on some recent work by Google, which we now describe.

Google Patent Embedding In 2019, Google Patent released an embedding representation of each publication available in their public dataset (hereafter, "GP embedding"). As detailed in Srebrovic (2019), embeddings are a set of techniques in natural language processing that map a text to a vector of real numbers. By leveraging methods such as neural networks, this mapping allow to significantly reduce the dimensionality of a text input.

The GP embedding is a vector of 64 dimensions that have been constructed in order to predict a patent's CPC (Cooperative Patent Code) from its text (including all metadata, abstract and body of the patent description). Each element of the vector is a continuous variable ranging between -1 and +1. It therefore summarize the text content of a vector in a simple algebraic representation which has the advantage of allowing to calculate the distance between two patents by taking the dot product between the two corresponding embeddings.

Formally, for each patent p, we let $\mathcal{E}(\mathbf{p})$ denote its embedding representation. We then define the distance between a patent p and a patent q as:

$$d(p,q) = \mathcal{E}(\mathbf{p}).\mathcal{E}(\mathbf{q})$$

Measure of novelty using text Using this distance measure, we can construct a novelty measure to capture radical innovation. The concept of novelty of a patent captures the extent to which a patent is significantly different from previous innovations in the same field. Typical measures of novelty look at the diversity of technological classes in the set of citing patents, or in the set of cited patents. These measures, sometimes called "originality" are limited by the fact that the average patent does not receive many citations.

Recently, the innovation literature has devoted much attention to using the text content of patent documents to refine some existing measures. For example, Kelly et al. (2018) shows how using the description of the innovation included in a patent publication can be used to build measures of similarity and novelty.

Here, we adapt their methodology. More precisely, we define novelty for each patent as the distance between its embedding and a reference point. This reference point is computed by calculating the unweighted average of all USPTO patents filed in the past 5 years and within the same technological class (we use 3-digit CPC classification, that is around 130 different categories). Formally, we define novelty $NOV(p)$ for each patent p as:

$$NOV(p) = \mathcal{E}(\mathbf{p})\frac{1}{N(k,t)}\sum_{q\in\mathcal{P}(k,t)}\mathcal{E}(\mathbf{q}) = \frac{1}{N(k,t)}\sum_{q\in\mathcal{P}(k,t)}d(p,q),$$

where k is the technological class of patent p,[39] $\mathcal{P}(k, t)$ is the set of USPTO patent filed between $t-5$ and $t-1$ and belonging to technological class k and $N(p, t)$ is its cardinal.

The static cross-section relationship between size and innovation when restricting attention to very novel patents (top 10%) and other patents respectively are shown in Figure 5(a). This graph is analogous to Figure 9. Likewise, results from regressions similar to that performed in Table 5 but using thresholds based on the value of novelty, are shown in Table D6.

Automation Patents that protect automation technologies have been the subject of a large body of work recently (see e.g. Dechezlepretre et al., 2020; Webb, 2019; Mann and Püttmann, 2018 for reviews). These papers typically use at the description (or abstract) of the patents to identify the occurrence of words that are usually associated with labor-saving technologies.

To build our automation measure, we use the work of Mann and Püttmann (2018) who look at the wording of USPTO patents and build a classifier to distinguish between automation and non-automation technologies. To apply their work to our set of patents, we once again leverage the GP embedding. Specifically, we regress the binary variable from Mann and Püttmann (2018) (1 if patent is classified as automation and 0 otherwise) on each of the 64 coordinates of the patent's embedding. We then use the estimated coefficients to predict the probability of being an automation patent for every patent owned by a French firm. Formally, we define our score of automation $A(p)$ for each patent p as:

$$A(p) = \sum_{i=1}^{64} \hat{\beta}_i \mathcal{E}(p)_i,$$

where $\hat{\beta}_i$ is the estimated coefficients from a model restricted to USPTO patents:

$$Y_q = \sum_{i=1}^{64} \beta_i \mathcal{E}(q)_i + \nu_t + \varepsilon_q,$$

for a patent q published during year t. In this model, Y_q is equal to 1 if the patent has been classified as an automation patent and ε is an error term.

[39] In the case where a patent has more than one CPC code, we consider patents from all the CPC codes in which case k represents the set of technological classes. In other words, we use a weighted average.

The underlying idea is that a linear combination of the embedding coordinates capture the feature included in the text that predict that a patent protects a labor-saving technology.

We show the cross section relationship between size and innovation result in Figure 5(b) and the dynamic regression in Table D7.

As an alternative to the automation measure, we also used a measure of the extent to which a patent protects a process innovation (as opposed to a product innovation) using the classification of Ashish Arora and Lee (2020). This uses the percentages of product or process related words in either the claims or the description of the patent publication document. It is likely the process innovations are more labor saving than product innovations, so the impact of regulation such fall more heavily on the product innovations. As with Mann and Püttmann (2018), this measure is only computed on USPTO patents, we leverage GP again, using the same methodology as the one described above to predict a corresponding value for our set of patents owned by French firms. We obtain broadly similar results. For example, splitting patents at the median level of "process", there are only signficant negative effects of the threshold on below median levels of process innovation, i.e. for product innovation.

Table D6: Regression results for different levels of the novelty of innovation

Novelty	Top 10%	Top 15%	Top 25%	Bottom 75%	Bottom 85%	Bottom 90%
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
$Shock_{t-2} \times L^{\star}_{t-2}$	-2.557	-2.963	-2.958*	-9.123***	-9.406***	-9.523***
	(1.580)	(1.973)	(1.634)	(3.108)	(2.783)	(3.009)
$L^{\star}_{t-2}$	0.040*	0.038	0.052	0.019	0.022	0.015
	(0.022)	(0.037)	(0.090)	(0.092)	(0.088)	(0.089)
$Shock_{t-2}$	-5.301	-7.666**	-8.231**	-4.802	-4.803	-7.035
	(3.190)	(3.644)	(3.324)	(4.024)	(4.527)	(4.663)
$log(L)_{t-2}$	0.002	0.010	-0.022	-0.035	-0.059	-0.057
	(0.011)	(0.011)	(0.020)	(0.030)	(0.035)	(0.033)
$Shock_{t-2} \times log(L)_{t-2}$	1.901*	2.776**	2.973**	2.034	2.094	2.800*
	(1.049)	(1.235)	(1.116)	(1.325)	(1.485)	(1.516)
Fixed Effects						
Sector×Year	✓	✓	✓	✓	✓	✓
Number Obs.	189,727	189,727	189,727	189,727	189,727	189,727

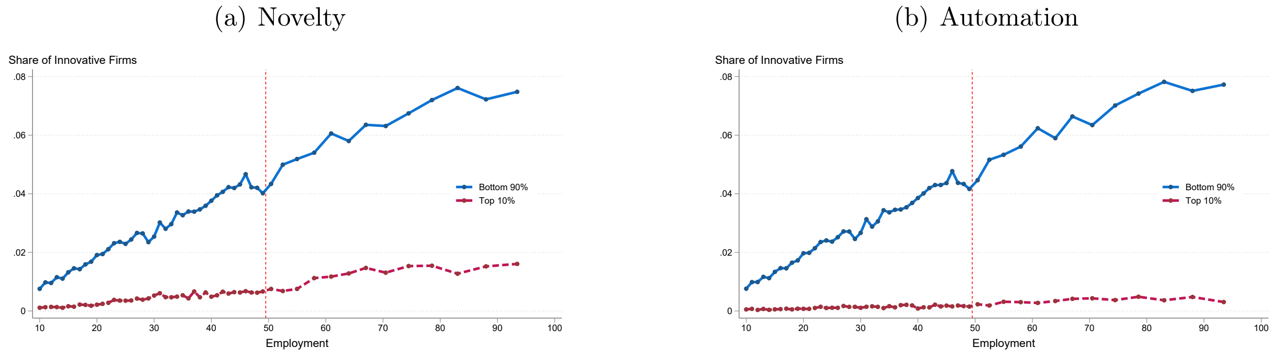
Notes: estimation results of the same model as in column 5 of Table 2. The dependent variable is the Davis and Haltiwanger (1992) growth rate in the number of priority patent applications between $t-1$ and t, restricting to the top 10% most novel (column 1), the top 15% most novel, the top 25% most novel (column 3), the bottom 85% most novel (column 4), the bottom 75% most novel (column 5) and the bottom 90% most novel (column 6). Definition of novelty is presented in Section 5.1.3. All models include a 2-digit NACE sector interacted with year fixed effects. Estimation period: 1997-2007. Standard errors are clustered at the 2-digit NACE sector level. ***, ** and * indicate p-value below 0.01, 0.05 and 0.1 respectively.

Table D7: Regression results for different levels of the automation

Auomation	Top 10% (1)	Top 15% (2)	Top 25% (3)	Bottom 75% (4)	Bottom 85% (5)	Bottom 90% (6)
$Shock_{t-2} \times L^{\star}_{t-2}$	0.640	-0.224	1.933	-10.695**	-10.272**	-10.828***
	(0.690)	(1.110)	(1.659)	(4.232)	(3.960)	(3.609)
$L^{\star}_{t-2}$	0.024	+0.03	-0.000	0.018	0.027	0.010
	(0.023)	(0.042)	(0.063)	(0.104)	(0.101)	(0.097)
$Shock_{t-2}$	0.203	-2.257**	0.277	-9.192*	-8.873*	-10.380**
	(1.095)	(0.925)	(2.461)	(4.495)	(4.452)	(4.320)
$log(L)_{t-2}$	-0.001	0.003	0.004	-0.051	-0.052	-0.051
	(0.005)	(0.005)	(0.008)	(0.034)	(0.036)	(0.038)
$Shock_{t-2} \times log(L)_{t-2}$	-0.206	0.699**	0.129	3.399**	3.526**	4.102***
	(0.409)	(0.335)	(0.834)	(1.530)	(1.519)	(1.418)
Fixed Effects						
Sector×Year	✓	✓	✓	✓	✓	✓
Number Obs.	189,727	189,727	189,727	189,727	189,727	189,727

Notes: estimation results of the same model as in column 5 of Table 2. The dependent variable is the Davis and Haltiwanger (1992) growth rate in the number of priority patent applications between $t-1$ and t, restricting to the top 10% patents that score highest in terms of predictive automation measure (column 1) and respectively top 15%, top 25%, bottom 25%, bottom 85% and bottom 90% patents. Definition of automation is presented in Section 5.2. All models include a 2-digit NACE sector interacted with year fixed effects. Estimation period: 1997-2007. Standard errors are clustered at the 2-digit NACE sector level. ***, ** and * indicate p-value below 0.01, 0.05 and 0.1 respectively.

Figure D5: Innovative firms at each employment level - novelty and automation



Notes: These Figures replicate Figure 9 but split patents between top 10% and bottom 90% according to their level of novelty (left panel) or their predicted level of automation (right panel).