Innovation and Top Income Inequality*

Philippe Aghion Ufuk Akcigit Antonin Bergeaud

Richard Blundell David Hémous

April 11, 2016

Abstract

In this paper we use cross-state panel and cross US commuting-zone data to look at the relationship between innovation, top income inequality and social mobility. We find positive and significant correlations between measures of innovation on the one hand, and top income inequality on the other hand. We also show that the correlations between innovation and broad measures of inequality are not significant, and that top income inequality is no longer correlated with highly lagged innovation. Next, using instrumentation analysis, we argue that these correlations at least partly reflect a causality from innovation to top income shares. Finally, we show that innovation, particularly by new entrants, is positively associated with social mobility, but less so in Metropolitan Statistical Areas with more intense lobbying activities.

JEL classification: O30, O31, O33, O34, O40, O43, O47, D63, J14, J15

Keywords: top income, inequality, innovation, patenting, citations, social mobility, incumbents, entrant.

*Addresses - Aghion: Harvard University, NBER and CIFAR. Akcigit: University of Chicago and NBER. Bergeaud: Banque de France. Blundell: University College London, Institute of Fiscal Studies, IZA and CEPR. Hémous: University of Zurich and CEPR. We are most grateful to John Van Reenen for detailed comments and advice throughout this project. We also thank Daron Acemoglu, Pierre Azoulay, Raj Chetty, Mathias Dewatripont, Peter Diamond, Thibault Fally, Maria Guadalupe, John Hassler, Elhanan Helpman, Chad Jones, Pete Klenow, Torsten Persson, Thomas Piketty, Andres Rodriguez-Clare, Emmanuel Saez, Stefanie Stantcheva, Scott Stern, Francesco Trebbi, Fabrizio Zilibotti, and seminar participants at MIT Sloan, INSEAD, the University of Zurich, Harvard University, The Paris School of Economics, Berkeley, the IIES at Stockholm University, Warwick University, Oxford, the London School of Economics, the IOG group at the Canadian Institute for Advanced Research, the NBER Summer Institute, and the 2016 ASSA meetings, for helpful comments and suggestions.

1 Introduction

That the past decades have witnessed a sharp increase in top income inequality worldwide and particularly in developed countries, is by now a widely acknowledged fact.[1] However no consensus has been reached as to the main underlying factors behind this surge in top income inequality. [2]In this paper we argue that, in a developed country like the US, innovation is certainly one such factor. For example, looking at the list of the wealthiest individuals across US states in 2015 compiled by Forbes (Brown, 2015), 11 out of 50 are listed as inventors in a US patent and many more manage or own firms that patent. More importantly, if we look at patenting and top income inequality in the US and other developed countries over the past decades, we see that these two variables tend to follow parallel evolution.

Thus Figure 1 below looks at patenting per 1000 inhabitants and the top 1% income share in the US since the 1960s: up to the early 1980s, both variables show essentially no trend but since then the two variables experience parallel upward trends.[3]

More closely related to our analysis in this paper, Figure 2 looks at the relationship between the increase in the log of innovation in a state between 1980 and 2005 (measured here by the number of citations within five years after patent application per inhabitant in the state) and the increase in the share of income held by the top 1% in that state over the same period. We see a clearly positive correlation between these two variables.[4] In this paper, we go further by using cross-state panel data to look at the relationship between top income inequality and innovation.

In a first part of the paper we develop a Schumpeterian growth model where growth results from quality-improving innovations that can be made in each sector either from the incumbent in the sector or from potential entrants. Facilitating innovation or entry increases the entrepreneurial share of income and spurs social mobility through creative destruction as employees' children more easily become business owners and vice versa. In particular, this model predicts that: (i) innovation by entrants and incumbents increases top income

[1]The worldwide interest for income and wealth inequality, has been spurred by popular books such as Goldin and Katz (2008), Deaton (2013) and Piketty (2014).

[2]Song *et al.* (2015) show that most of the rise in earnings inequality can be explained by the rise in across-firm inequality rather than within-firm inequality.

[3]The figures in this introduction use unweighted patent counts as measure of innovation. Using citation-weighted patent counts yields similar patterns, although the series for unweighted patent counts are available over a longer period.

[4]This does not mean that all top 1% income earners are inventors or that innovation only increases the income of inventors. Indeed Table 6a from Bakija et al. (2008) shows an 11.2 point growth of the top 1% in the US as a whole between 1979 and 2005, but only a 1.37 point out of the 11.2 is accounted for by entrepreneurs, technical occupations, scientists and business operations. The bulk of the growth in the top 1% accrues to financiers, lawyers and executive managers some of whom typically accompany and benefit from the innovation process.

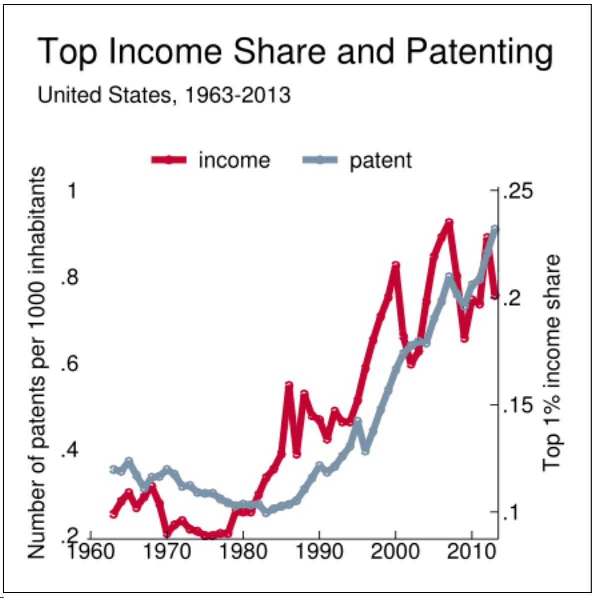

Figure 1: This figure plots the number of patent applications per 1000 inhabitant against the top 1% income share for the USA as a whole. Observations span the years 1963-2013.

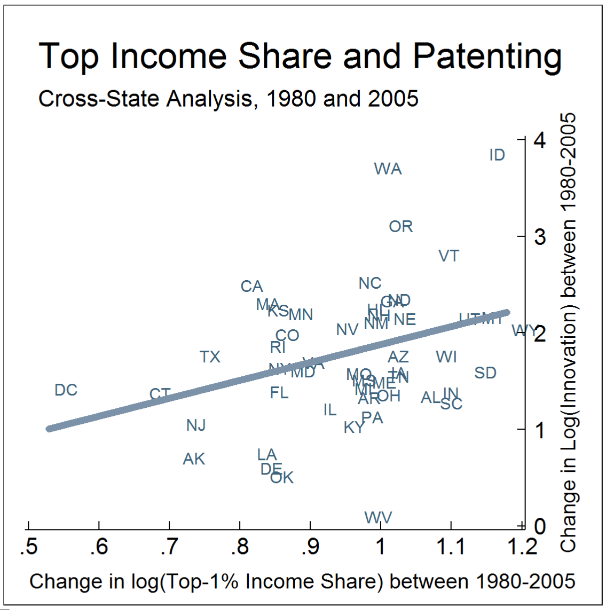

Figure 2: This figure plots the difference of the log of the number of citations per capita against the difference of the log of the top 1% income share in 1980 and 2005. Observations are computed at the US state level.

inequality; (ii) innovation by entrants increases social mobility; (iii) entry barriers lower the positive effects of entrants' innovations on top income inequality and social mobility. In the remaining part of the paper, we confront these predictions with available cross state panel and cross commuting zone data.

We then start our empirical analysis by exploring correlations between innovation and various measures of inequality using OLS regressions. Our main findings can be summarized as follows. First, the top 1% income share in a given US state in a given year, is positively and significantly correlated with the state's degree of innovation, measured either by the flow of patents or by the quality-adjusted amount of innovation in this state in that year, as reflected by citations. Second, we find that innovation is less positively or even negatively correlated with measures of inequality which do not emphasize the very top incomes, in particular the top 2 to 10% income shares (i.e. excluding the top 1%), or broader measures of inequality like the Gini coefficient, as suggested by Figure 3 below.[5] Next, looking at the relationship between inequality and innovation at various lags, we find that the correlation

[5]Figure 3 plots the average top-1% income share and the bottom 99% Gini index as a function of their corresponding innovation percentiles. The bottom 99% Gini is the Gini coefficient when the top 1% of the income distribution is removed. Innovation percentiles are computed using the US state-year pairs from 1975 to 2010. Each series is normalized by its value in the lowest innovation percentile.

between innovation and the top 1% income share is temporary. Finally, we find that the correlation between innovation and top income inequality is dampened in states with higher lobbying intensity.

Next, we argue that the correlation between innovation and top inequality at least partly reflects a causal effect of innovation-led growth on top incomes. We instrument for innovation using data on the appropriation committees of the Senate (following Aghion *et al.*, 2009). We find that all the broad OLS results in Section 4 are confirmed by the corresponding IV regressions.

Our results pass a number of robustness tests. First, we add a second instrument for innovation in each state which relies on knowledge spillovers from the other states. We show that when the two instruments are used jointly, the overidentification test does not reject the null hypothesis that the instruments are uncorrelated with the error term. In other words, we do not reject the validity of the instruments. Second, we show that the positive and significant correlation between innovation and top income shares in cross state panel regressions, is robust to introducing various proxies reflecting the importance of the financial sector, to including top marginal tax rates as control variables (whether on capital, labor or interest income), and to controlling for sectors' size or for potential agglomeration effects.

Finally, when looking at the relationship between innovation and social mobility, using cross-section regressions performed at the commuting zone (CZ) level, we find that: (i) innovation is positively correlated with upward social mobility (Figure 4 below[6]); (ii) the positive correlation between innovation and social mobility, is driven mainly by entrant innovators and less so by incumbent innovators, and it is dampened in MSAs with higher lobbying intensity.

The analysis in this paper relates to several strands of literature. First, to the endogenous growth literature (Romer, 1990; Aghion and Howitt, 1992). We contribute to this literature, first by introducing social mobility into the picture and linking it to creative destruction, and second by looking explicitly at the effects of innovation on top income shares.[7]

Second, our paper relates to an empirical literature on inequality and growth. Most

[6]Figure 4 plots the logarithm of the number of patent applications per capita (x-axis) against the logarithm of social mobility (y-axis). Social mobility is computed as the probability to belong to the highest quintile of the income distribution in 2010 (when aged circa 30) when parents belonged to the lowest quintile in 1996 (when aged circa 16). Observations are computed at the Commuting Zones level (569 observations). The number of patents is averaged from 2006 to 2010.

[7]Hassler and Rodriguez-Mora (2000) analyze the relationship between growth and intergenerational mobility in a model which may feature multiple equilibria, some with high growth and high social mobility and others with low growth and low social mobility. Multiple equilibria arise because in a high growth environment, inherited knowledge depreciates faster, which reduces the advantage of incumbents. In that paper however, growth is driven by externalities instead of resulting from innovations.

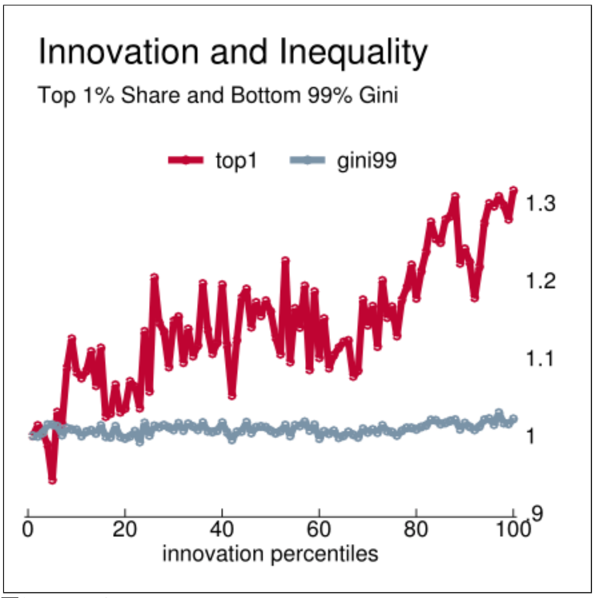

Figure 3: SEE FOOTNOTE 5 FOR EXPLANATIONS.

Figure 4: SEE FOOTNOTE 6 FOR EXPLANATIONS.

closely related to our analysis, Frank (2009) finds a positive relationship between both the top 10% and top 1% income shares and growth across US states; however, he does not establish any causal link from growth to top income inequality, nor does he consider innovation or social mobility.[8]

Third, a large literature on skill-biased technical change aims at explaining the increase in labor income inequality since the 1970's.[9] While this literature focuses on the *direction* of innovation and on broad measures of labor income inequality (such as the skill-premium), our paper is more directly concerned with the rise of the top 1% and how it relates with the *rate* and *quality* of innovation (in fact our results suggest that innovation does not have a strong impact on broad measures of inequality compared to their impact on top income shares).

Fourth, our focus on top incomes links our paper to a large literature documenting a sharp increase in top income inequality over the past decades (in particular, see Piketty

[8]Acemoglu and Robinson (2015) also reports a positive correlation between top income inequality and growth in panel data at the country level (or at least no evidence of a negative correlation).

[9]In particular, Katz and Murphy (1992) and Goldin and Katz (2008) have shown that technical change has been skill-biased in the 20th century. Acemoglu (1998, 2002 and 2007) sees the skill distribution as determining the direction of technological change, while Hémous and Olsen (2014) argue that the incentive to automate low-skill tasks naturally increases as an economy develops. Several papers (Aghion and Howitt, 1997; Caselli, 1999; Galor and Moav, 2000) see General Purpose Technologies (GPT) as lying behind the increase in inequality, as the arrival of a GPT favors workers who adapt faster to the detriment of the rest of the population. Krusell, Ohanian, Ríos-Rull and Violante (2000) show how with capital-skill complementarity, the increase in the equipment stock can account for the increase in the skill premium.

and Saez, 2003). We contribute to this line of research by arguing that increases in top 1% income shares, are at least in part caused by increases in innovation-led growth.[10]

Fiflth, the part of our analysis on social mobility and innovation, directly builds on Chetty *et al.* (**?**) who collect information on intergenerational mobility across US Commuting Zones using tax data on parents and children.[11] We contribute to this line of research by linking social mobility to innovation and creative destruction.

Most closely related to our paper is Jones and Kim (2014), who also develop a Schumpeterian model to explain the dynamics of top income inequality. In their model, growth results from both, the accumulation of experience or knowledge by incumbents (which may in turn result from incumbent innovation) and creative destruction by entrants. The former increases top income inequality whereas the latter reduces it by allowing entrants to catch up with incumbents.[12] In our model instead, a new (entrant) innovation increases mark-ups in the corresponding sector, whereas in the absence of a new innovation, mark-ups are partly eroded as a result of imitation. On the other hand, the two papers have in common the ideas: (i) that innovation and creative destruction are key factors in the dynamics of top income inequality; (ii) that fostering entrant innovation contributes to making growth more "inclusive".[13]

The remaining part of the paper is organized as follows. Section 2 outlays a simple Schumpeterian model to guide our analysis of the relationship between innovation-led growth, top incomes, and social mobility. Section 3 presents our cross-state panel data and our measures of inequality and innovation. Section 4 presents our OLS regression results. Section 5 presents our IV results. Section 6 performs robustness tests. Section 7 looks at the relationship between innovation and social mobility. And Section 8 concludes.

The main tables (Table 1 to Table 16) are displayed at the end of the main text. The Online Appendix A contains the theoretical proofs. And the Online Appendix B displays

[10]Rosen (1981) emphasizes the link between the rise of superstars and market integration: namely, as markets become more integrated, more productive firms can capture a larger income share, which translates into higher income for its owners and managers. Similarly, Gabaix and Landier (2008) show that the increase in the size of some firms can account for the increase in their CEO's pay. Our analysis is consistent with this line of work, to the extent that successful innovation is a main factor driving differences in productivities across firms, and therefore in firms' size.

[11]For prior surveys on intergenerational mobility, see Solon (1999) and Black and Devereux (2011).

[12]More specifically, in Jones and Kim (2014) entrants innovation only reduces income inequality because it affects incumbents' efforts. Therefore in their model an exogenous increase in entrant innovation will not affect inequality if it is not anticipated by incumbents.

[13]Indeed, we show that entrant innovation is positively associated with social mobility. Moreover, if, as we shall see below, incumbent innovation and entrant innovation contribute to a comparable extent to increasing the top 1% income share, additional regressions shown in Appendix (see Table B1) suggest that incumbent innovation contributes more to increasing the top 0.1% share than entrant innovation (and even more for the top 0.01% share).

the additional tables (Tables B1 to B12).

2 Theory

In this section we develop a simple Schumpeterian growth model to explain why increased R&D productivity increases both the top income share and social mobility.

2.1 Baseline model

Consider the following discrete time model. The economy is populated by a continuum of individuals. At any point in time, there is a measure $L + 1$ of individuals in the economy, a mass 1 are capital owners who own the firms and the rest of the population works as production workers (with $L \geq 1$). Each individual lives only for one period. Every period, a new generation of individuals is born and individuals that are born to current firm owners inherit the firm from their parents. The rest of the population works in production unless they successfully innovate and replace incumbents' children.

2.1.1 Production

A final good is produced according to the following Cobb-Douglas technology:

$$\ln Y_t = \int_0^1 \ln y_{it} di, \tag{1}$$

where y_{it} is the amount of intermediate input i used for final production at date t. Each intermediate is produced with a linear production function

$$y_{it} = q_{it} l_{it}, \tag{2}$$

where l_{it} is the amount of labor used to produce intermediate input i at date t, and q_{it} is labor productivity. Each intermediate i is produced by a monopolist who faces a competitive fringe from the previous technology in that sector.

2.1.2 Innovation

Whenever there is a new innovation in any sector i in period t, quality in that sector improves by a multiplicative term $\eta_H > 1$ so that:

$$q_{i,t} = \eta_H q_{i,t-1}.$$

In the meantime, the previous technological vintage $q_{i,t-1}$ becomes publicly available, so that the innovator in sector i obtains a technological lead of η_H over potential competitors.

At the end of period t, other firms can partly imitate the (incumbent) innovator's technology so that, in the absence of a new innovation in period $t+1$, the technological lead enjoyed by the incumbent firm in sector i shrinks to η_L with $1 < \eta_L < \eta_H$.

Overall, the technological lead enjoyed by the incumbent producer in any sector i takes two values: η_H in periods with innovation and $\eta_L < \eta_H$ in periods without innovation.[14]

Finally, we assume that an incumbent producer that has not recently innovated, can still resort to lobbying in order to prevent entry by an outside innovator. Lobbying is successful with exogenous probability z, in which case, the innovation is not implemented, and the incumbent remains the technological leader in the sector (with a lead equal to η_L).

Both potential new entrants and incumbents have access to the following innovation technology. By spending

$$C_{K,t}(x) = \theta_K \frac{x^2}{2} Y_t$$

an incumbent ($K = I$) or entrant ($K = E$) can innovate with probability x. A reduction in θ_K captures an increase in R&D productivity or R&D support, and we allow for it to differ between entrants and incumbents.

2.1.3 Timing of events

Each period unfolds as follows:

1. In each line i where an innovation occurred in the previous period, followers copy the corresponding technology so that the technological lead of the incumbent shrinks to η_L.

2. In each line i, a single potential entrant is drawn from the mass of workers' offsprings and spends $C_{E,t}(x_{E,i})$ and the offspring of the incumbent in sector i spends $C_{I,t}(x_{I,i})$.

3. With probability $(1-z)\,x_{E,i}$ the entrant succeeds, replaces the incumbent and obtains a technological lead η_H, with probability $x_{I,i}$ the incumbent succeeds and improves its technological lead from η_L to η_H, with probability $1-(1-z)\,x_{E,i} - x_{I,i}$, there is no successful innovation and the incumbent stays the leader with a technological lead of η_L.[15]

[14] The details of the imitation-innovation sequence do not matter for our results, what matters is that innovation increases the technological lead of the incumbent producer over its competitive fringe.

[15] For simplicity, we rule out the possibility that both agents innovate in the same period, so that in a given

4. Production and consumption take place and the period ends.

2.2 Solving the model

We solve the model in two steps: first, we compute the income shares of entrepreneurs and workers and the rate of upward social mobility (from being a worker to becoming an entrepreneur) for given innovation rates by entrants and incumbents; second, we endogeneize the entrants' and incumbents' innovation rates.

2.2.1 Income shares and social mobility for given innovation rates

In this subsection we assume that in all sectors, potential entrants innovate at some exogenous rate x_{Et} and incumbents innovate at some exogenous rate x_{It} at date t.

Using (2), the marginal cost of production of (the leading) intermediate producer i at time t is

$$MC_{it} = \frac{w_t}{q_{i,t}}.$$

Since the leader and the fringe enter Bertrand competition, the price charged at time t by intermediate producer i is simply a mark-up over the marginal cost equal to the size of the technological lead, i.e.

$$p_{i,t} = \frac{w_t \eta_{it}}{q_{i,t}}, \tag{3}$$

where $\eta_{i,t} \in \{\eta_H, \eta_L\}$. Therefore innovating allows the technological leader to charge temporarily a higher mark-up.

Using the fact that the final good sector spends the same amount Y_t on all intermediate goods (a consequence of the Cobb-Douglas technology assumption), we have in equilibrium:

$$p_{i,t} y_{it} = Y_t \text{ for all } i. \tag{4}$$

This, together with (3) and (2), allows us to immediately express the labor demand and the equilibrium profit in any sector i at date t.

Labor demand by producer i at time t is given by:

$$l_{it} = \frac{Y_t}{w_t \eta_{it}}.$$

sector, innovations by the incumbent and the entrant are not independent events. This can be microfounded in the following way. Assume that every period there is a mass 1 of ideas, and only one idea is succesful. Research efforts x_E and x_I represent the mass of ideas that a firm investigates. Firms can observe each other actions, therefore in equilibrium they will never choose to look for the same idea provided that $x_E^* + x_I^* < 1$, which is satisfied for θ_K sufficiently large.

Equilibrium profits in sector i at time t are equal to:

$$\Pi_{it} = (p_{it} - MC_{it})y_{it} = \frac{\eta_{it} - 1}{\eta_{it}} Y_t.$$

Hence profits are higher if the incumbent has recently innovated, namely:

$$\Pi_{H,t} = \underbrace{\frac{\eta_H - 1}{\eta_H}}_{\equiv \pi_H} Y_t > \Pi_{L,t} = \underbrace{\frac{\eta_L - 1}{\eta_L}}_{\equiv \pi_L} Y_t.$$

We can now derive the expressions for the income shares of workers and entrepreneurs and for the rate of upward social mobility. Let μ_t denote the fraction of high-mark-up sectors (i.e. with $\eta_{it} = \eta_H$) at date t. Labor market clearing at date t implies that:

$$L = \int_0^1 l_{it} di = \int_0^1 \frac{Y_t}{w_t \eta_{it}} di = \frac{Y_t}{w_t} \left[\frac{\mu_t}{\eta_H} + \frac{1 - \mu_t}{\eta_L} \right]$$

We restrict attention to the case where $\eta_L - 1 > 1/L$, which ensures that regardless of the equilibrium value of μ_t,

$$w_t < \Pi_{L,t},$$

so that top incomes are earned by entrepreneurs. As a result, the entrepreneur share of income is a proxy for top income inequality (defined as the share of income that goes to the top earners—not as inequality within top-earners).

Hence the share of income earned by workers (wage share) at time t is equal to:

$$wages\_share_t = \frac{w_t L}{Y_t} = \frac{\mu_t}{\eta_H} + \frac{1 - \mu_t}{\eta_L}. \tag{5}$$

whereas the gross share of income earned by entrepreneurs (entrepreneurs share) at time t is equal to:

$$entrepreneur\_share_t = \frac{\mu_t \Pi_{H,t} + (1 - \mu_t) \Pi_{L,t}}{Y_t} = 1 - \frac{\mu_t}{\eta_H} - \frac{1 - \mu_t}{\eta_L}. \tag{6}$$

This entrepreneur share is "gross" in the sense that it does not take into account any potential monetary costs of innovation (and similarly all our share measures are expressed as functions of total output and not of net income—see below for the net shares).

Since mark-ups are larger in sectors with new technologies, aggregate income shifts from workers to entrepreneurs in relative terms whenever the equilibrium fraction of product lines with new technologies μ_t increases. But by the law of large numbers this fraction is equal to the probability of an innovation by either the incumbent or a potential entrant in any intermediate good sector.

More formally, we have:

$$\mu_t = x_{It} + (1 - z)\, x_{Et}, \tag{7}$$

which increases with the innovation intensities of both incumbents and entrants, but to a lesser extent with respect to entrants' innovations the higher the entry barriers z are.

Finally, we measure upward social mobility by the probability Ψ_t that the offspring of a worker becomes a business owner. This in turn happens only if this individual gets to be a potential entrant and then manages to innovate and to avoid the entry barrier; therefore

$$\Psi_t = x_{Et}\,(1 - z)\,/L, \tag{8}$$

which is increasing in entrant's innovation intensity x_{Et} but less so the higher the entry barriers z are. This yields:

Proposition 1 *(i) A higher rate of innovation by a potential entrant, x_{Et}, is associated with a higher entrepreneur share of income and a higher rate of social mobility, but less so the higher the entry barriers z are; (ii) A higher rate of innovation by an incumbent, x_{It}, is associated with a higher entrepreneur share of income but has no direct impact on social mobility.*

Remark: That the equilibrium *share* of wage income in total income decreases with the fraction of high mark-up sectors μ_t, and therefore with the innovation intensities of entrants and incumbents, does not imply that the equilibrium *level* of wages also declines. In fact the opposite occurs.[16] In addition, note that the entrepreneurial share is independent of innovation intensities in previous periods. Therefore, a temporary increase in current

[16] To see this more formally, we can compute the equilibrium level of wages by plugging (4) and (3) in (1), which yields:

$$w_t = Q_t / \left(\eta_H^{\mu_t} \eta_L^{1-\mu_t}\right), \tag{9}$$

where Q_t is the quality index defined as $Q_t = \exp \int_0^1 \ln q_{it} di$. The law of motion for the quality index is computed as

$$Q_t = \exp \int_0^1 \left[\mu_t \ln \eta_H q_{it-1} + (1 - \mu_t) \ln q_{it-1}\right] di = Q_{t-1} \eta_H^{\mu_t}. \tag{10}$$

Therefore, for given technology level at time $t - 1$, the equilibrium wage is given by

$$w_t = \eta_L^{\mu_t - 1} Q_{t-1}.$$

This last equation shows that the overall effect of an increase in innovation intensities is to increase the contemporaneous equilibrium wage, even though it also shifts some income share towards entrepreneurs.

innovation only leads to a temporary increase in the entrepreneurial share: once imitation occurs, the gains from the current burst in innovation will be equally shared by workers and entrepreneurs.

2.2.2 Endogenous innovation

We now turn to the endogenous determination of the innovation rates of entrants and incumbents. The offspring of the previous period's incumbent solves the following maximization problem:

$$\max_{x_I} \left\{ x_I \pi_H Y_t + (1 - x_I - (1-z)\, x_E^*)\, \pi_L Y_t + (1-z)\, x_E^* w_t - \theta_I \frac{x_I^2}{2} Y_t \right\}.$$

This expression states that the offspring of an incumbent can already collect the profits of the firm that she inherited ($\pi_L Y_t$), but also has the chance of making higher profit ($\pi_H Y_t$) by innovating with probability x_I. Clearly the optimal innovation decision is simply

$$x_{I,t} = x_I^* = \frac{\pi_H - \pi_L}{\theta_I} = \left(\frac{1}{\eta_L} - \frac{1}{\eta_H} \right) \frac{1}{\theta_I}, \tag{11}$$

which decreases with incumbent R&D cost parameter θ_I.

A potential entrant in sector i solves the following maximization problem:

$$\max_{x_E} \left\{ (1-z)\, x_E \pi_H Y_t + (1 - x_E (1-z))\, w_t - \theta_E \frac{x_E^2}{2} Y_t \right\},$$

since a new entrant chooses its innovation rate with the outside option being a production worker who receives wage w_t. Using equation (5), taking first order conditions, and using our assumption that $w_t < \pi_L Y_t$, we can express the entrant innovation rate as

$$x_{E,t} = x_E^* = \left(\pi_H - \frac{1}{L} \left[\frac{\mu_t}{\eta_H} + \frac{1-\mu_t}{\eta_L} \right] \right) \frac{(1-z)}{\theta_E}, \tag{12}$$

which implies that entrants innovate in equilibrium since $\pi_H > \pi_L > w/Y$.

Since in equilibrium $\mu^* = x_I^* + (1-z)\, x_E^*$, the equilibrium innovation rate for entrants is simply given by

$$x_E^* = \frac{\left(\pi_H - \frac{1}{L}\frac{1}{\eta_L} + \frac{1}{L} \left(\frac{1}{\eta_L} - \frac{1}{\eta_H} \right) x_I^* \right) (1-z)}{\theta_E - \frac{1}{L} (1-z)^2 \left(\frac{1}{\eta_L} - \frac{1}{\eta_H} \right)}. \tag{13}$$

Throughout this section, we implicitly assume that θ_I and θ_E are sufficiently large that

$x_E^* + x_I^* < 1$.

Therefore lower barriers to entry (i.e. a lower z) and less costly R&D for entrants (lower θ_E) both increase the entrants' innovation rate (as $1/\eta_L - 1/\eta_H > 0$). Less costly incumbent R&D also increases the entrant innovation rate since x_I^* is decreasing in θ_I.[17]

Intuitively, high mark-up sectors are those where an innovation just occurred and was not blocked, so a reduction in either entrants' or incumbents' R&D costs increases the share of high mark-up sectors in the economy and thereby the gross entrepreneurs' share of income. To the extent that higher entry barriers dampen the positive correlation between the entrants' innovation rate and the entrepreneurial share of income, they will also dampen the positive effects of a reduction in entrants' or incumbents' R&D costs on the entrepreneurial share of income.

Finally, equation (8) immediately implies that a reduction in entrants' or incumbents' R&D costs increases social mobility but less so the higher the barriers to entry are. We have thus established (proof in Appendix A.1):

Proposition 2 *An increase in R&D productivity (whether it is associated with a reduction in θ_I or in θ_E), leads to an increase in the innovation rates x_I^* and x_E^* but less so the higher the entry barriers z are; consequently, it leads to higher growth, higher entrepreneur share and higher social mobility but less so the higher the entry barriers are.*

2.2.3 Entrepreneurial share of income net of innovation costs

So far we computed gross shares of income, ignoring innovation expenditures.[18] If we now discount these expenditures, the ratio between net entrepreneurial income and labor income can be written as:

$$\begin{aligned} rel\_net\_share &= \left(Entrepreneur\_share_t - \theta_E \frac{x_E^2}{2} - \theta_I \frac{x_I^2}{2} \right) / \left(\frac{w_t}{Y_t} L \right) \\ &= \left(\pi_L + \frac{\pi_H - \pi_L}{2} x_I^* + \left(\frac{\pi_H}{2} + \frac{w_t}{2Y_t} - \pi_L \right) (1-z)\, x_E^* \right) / \left(\frac{w_t}{Y_t} L \right) \quad (14) \end{aligned}$$

where we used (6), (7) and the equilibrium values (11) and (12). This expression shows that a higher rate of incumbent innovation will raise the net entrepreneur share of income,

[17] x_E^* increases with x_I^* because more innovation by incumbents lowers the equilibrium wage which decreases the opportunity cost of innovation for an entrant. This general equilibrium effect rests on the assumption that incumbents and entrants cannot both innovate in the same period.

[18] Not factoring innovation costs in our computation of entrepreneur shares of incomne amounts to treating those as private utility costs. Also in practice entrepreneurial incomes are typically generated after the innovation costs are sunk, even though in our model we assume that innovation expenditures and entrepreneurial incomes occur within the same period.

whereas a higher rate of entrant innovation will only raise the net entrepreneurial share of income if $\frac{1}{2}\pi_H + \frac{1}{2}\frac{w_t}{Y_t} - \pi_L > 0$ (which occurs in particular if $\pi_H > 2\pi_L$). This in turn relates to the creative destruction nature of entrant's innovation: a successful entrant gains $\pi_H Y_t - w_t$ by innovating but she destroys the rents $\pi_L Y_t$ of the incumbent. Formally, we can show (see Appendix A.1):

Proposition 3 *An increase in incumbent R&D productivity (lower θ_I) leads to an increase in the relative shares of net entrepreneurial income over labor income. An increase in entrant R&D productivity (lower θ_E) also leads to an increase in the relative shares of net entrepreneurial income over labor income whenever $\frac{1}{2}\pi_H + \frac{1}{2}\frac{w_t}{Y_t} - \pi_L > 0$.*

On the other hand, we find that when L is large and π_H is close enough to π_L, then an increase in the productivity of entrant R&D will shift income towards workers instead of entrepreneurs, and therefore will contribute to a reduction in inequality. This result is in the vein of Jones and Kim (2014).

2.2.4 Impact of mark-ups on innovation and inequality

Our discussion so far pointed to a causality from innovation to top income inequality and social mobility. However the model also speaks to the reverse causality from top inequality to innovation. First, a higher innovation size η_H leads to a higher mark-up for firms which have successfully innovated. As a result, it increases the entrepreneur share for given innovation rate (see (6)). Meanwhile a higher η_H increases incumbents' (11) and (13) entrants' innovation rates, which further increases the entrepreneur share of income.

More interestingly perhaps, a higher η_L increases the mark-up of non-innovators, and thereby increases the entrepreneur share for a given innovation rate (see (6) and recall that $(1-z)\,x^* + \tilde{x}^* < 1$). Yet, it decreases incumbents' innovation rate since their net reward from innovation is lower. In the special case where $\theta_I = \theta_E$ this leads to a decrease in the total innovation rate (see Appendix A.2). For a sufficiently high R&D cost (θ high), the overall impact on the entrepreneur share remains positive. Therefore a higher η_L can contribute to a negative correlation between innovation and the entrepreneur share.

2.2.5 Shared rents from innovation

In the model so far, all the rents from innovation accrue to an individual entrepreneur who fully owns her firm. In reality though, the returns from innovation are shared among several actors (inventors, developers, the firm's CEO, financiers,...—see Aghion and Tirole, 1994, for a theoretical model of the relationship between inventors and developers and financiers of an innovation; Hall *et al.* (2005) show empirically that innovation increases firm value; and Balkin *et al.* (2001) show that innovation increases CEO's pay in high-technology firms). We show this formally in Appendix A.3 where we extend our analysis, first to the case where the innovation process involves an inventor and a CEO, second to the case where the inventor is distinct from the firm's owner(s). Our theoretical results are robust to these extensions.

2.3 Predictions

We can summarize the main predictions from the above theoretical discussion as follows.

- Innovation by both entrants and incumbents, increases top income inequality;
- Innovation by entrants increases social mobility;
- Entry barriers lower the positive effect of entrants' innovation on top income inequality and on social mobility.

Before we confront these predictions to the data, note that the above model also predicts that national income shifts away from labor towards firm owners as innovation intensifies. This is in line with findings from the recent literature on declining labor share (e.g. see Elsby *et al.* 2013 and Karabarbounis and Neiman 2014). In fact Figures 5 and 6 show that over the past forty years in the US, the profit share increased and the labor share decreased (one minus the labor share increased) in ways that paralleled the acceleration in innovation. This provides additional support for our model.

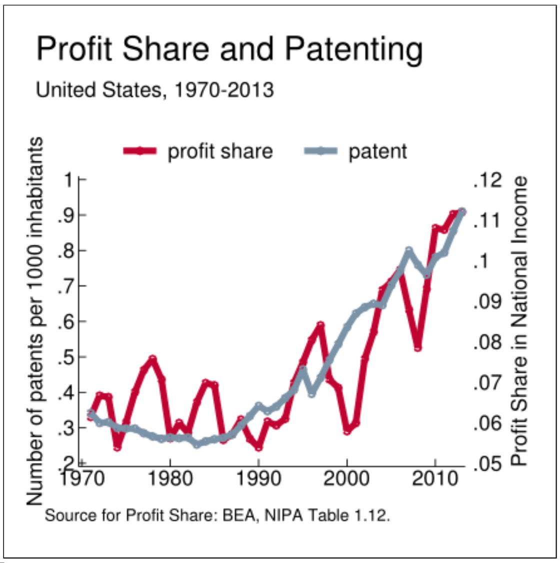

Figure 5: Profit Share in National Income

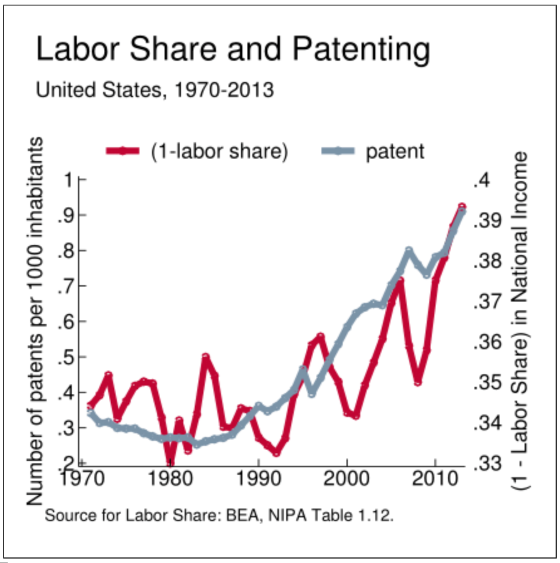

Figure 6: Labor Share in National Income

3 The empirical framework

In this section we present our measures of inequality and innovation and the databases used to compute these measures. Then we describe our estimation strategy.

3.1 Data and measurement

Our core empirical analysis is carried out at the US state level. Our dataset starts in 1975, a time range imposed upon us by the availability of patent data.

3.1.1 Inequality

The data on the share of income owned by the top 1% of the income distribution for our cross-US-state panel analysis, are drawn from the US State-Level Income Inequality Database (Frank, 2009, updated in 2015). From the same data source, we also gather information on alternative measures of inequality: namely, the top 0.01, 0.1, 0.5, 5 and 10% income shares, the Atkinson Index (with a coefficient of 0.5), the Theil Index and the Gini Index. These data are available from 1916 to 2013 but we restrict attention to the period after 1975. We end up with a balanced panel of 51 states (we include Alaska and Hawaii and count the District of Columbia as a "state") over a maximum time period of 36 years. In 2013, the three states with the highest share of total income earned by the richest 1% are New-York, Connecticut, and Wyoming with respectively 31.8%, 30.8% and 29.6% whereas Iowa, Hawaii and Alaska are the states with the lowest share earned by the top 1% (respectively 11.7%, 11.4% and 11.1%). In every US state, the top 1% income share has increased between 1975 and 2013, the unweighted mean value was around 8.4% in 1975 and reached 20.4% in 2007 before slowly decreasing to 17.1% in 2013. In addition, the heterogeneity in top income shares across states is larger in the recent period than it was during the 1970s, with a cross-state coefficient of variation multiplied by 2.2 between 1975 and 2013. The states that experienced the fastest growth in the top 1% income share during the considered time period are Wyoming,Idaho, Montana and South Dakota; on the other hand DC, Connecticut, New Jersey and Arkansas experienced the lowest growth in that share.

Note that the US State-Level Income Inequality Database provides information on the adjusted gross income from the IRS. This is a broad measure of pre-tax (and pre-transfer) income which includes wages, entrepreneurial income and capital income (including realized capital gains). Unfortunately it is not possible to decompose total income in the various sources of income (wage, entrepreneurial or capital incomes) with this dataset. In contrast, the World Top Income Database (Alvaredo *et al.* 2014), allows us to assess the composition of the top 1% income share. On average between 1975 and 2013, wage income represented 59.3% and entrepreneurial income 22.8% of the total income earned by the top 1%, while for the top 10%, wage income represented 76.9% and entrepreneurial income 12.9% of total income. In our baseline model, entrepreneurs are those directly benefiting from innovation. In practice, innovation benefits are shared between firm owners, top managers and inventors, thus innovation affects all sources of income within the top 1% (as highlighted in Appendix A.3). Yet, the fact that entrepreneurial income is over-represented in the top 1% income relative to wage income, suggests that our baseline model captures an important aspect in the evolution of top income inequality.

3.1.2 Innovation

When looking at cross state or more local levels, the US patent office (USPTO) provides complete statistics for patents granted between the years 1975 and 2014. For each patent, it provides information on the state of residence of the patent *inventor*, the date of application of the patent and a link to every citing patents granted before 2014. This citation network between patents enables us to construct several estimates for the quality of innovation as described below. Since a patent can be associated with more than one inventor and since coauthors of a given patent do not necessarily live in the same state, we assume that patents are split evenly between inventors and thus we attribute only a fraction of the patent to each inventor. A patent is also associated with an *assignee* that owns the right to the patent. Usually, the assignee is the firm employing the inventor, and for independent inventors the assignee and the inventor are the same person. We chose to locate each patent according to the US state where its inventor lives and works. Although the inventor's location might occasionally differ from the assignee's location, most of the time the two locations coincide (the correlation between the two is above 92%).[19] Finally, in line with the patenting literature, we focus on "utility patents" which cover 90% of all patents at the USPTO.[20]

We associate a patent with its year of application which corresponds to the year when the provisional application is considered to be complete by the USPTO and a filing date is set. However, we only consider patents that were ultimately granted by 2014. For that reason,

[19] For example, Delaware and DC are states for which the inventor's address is more likely to differ from the assignee's address for fiscal reasons.

[20] The USPTO classification considers three types of patents according to the official documentation: utility patents that are used to protect a new and useful invention, for example a new machine, or an improvement to an existing process; design patents that are used to protect a new design of a manufactured object; and plant patents that protect some new varieties of plants. Among those three types of patents, the first is presumably the best proxy for innovation, and it is the only type of patents for which we have complete data.

our data suffer from a truncation bias due to the time lag between application and grant. The USPTO considered in the end of 2012 that a patent application should be considered to be 95% complete for applications filed in 2004.[21] By the same logic, we consider that by the end of 2014, our patent data are essentially complete up to 2006. For the remaining years between 2006 and 2009, we correct for truncation bias using the distribution of time lags between the application and granting dates to extrapolate the number of patents by states following Hall *et al.* (2001). The small number of observed patents after 2009 led us to stop the correction in that year.

Simply counting the number of patents granted by their application date is a crude measure of innovation as it does not differentiate between a patent that made a significant contribution to science and a more incremental one. The USPTO database, provides sufficiently exhaustive information on patent citation to compute indicators which better measure the quality of innovation. We consider five measures of innovation quality.

- *5-year window citations counter*: this variable measures the number of citations received within no more than 5 years after the application date. This number has been corrected to account for different propensity to cite across sectors and across time. In addition, because of the drop in the number of observed completed patents in the patent data after 2006, we need to correct for the truncation bias in citations. We did so by following Hall et al. (2001). We consider that the 5-year citation counter series is reliable up to 2006.

- *Is the patent among the 5% (resp. 1%) most cited in the year* according to the previous measure? This is a dummy variable equal to one if the patent applied for in a given year belong to the top 5% (resp. 1%) most cited patents in the next five years following its publication. Because this measure is based on the number of citations within a 5-year window, the corresponding series is stopped in 2006. A rational for using this measure, as argued in Abrams *et al.* (2013), has to do with the existence of potential non-linearities between the value of a patent and the number of forward citations.

- *Patent breadth*, defined as the number of claims in a patent. As argued in Akcigit *et al.* (2015), it is common to use patent claims to proxy for patent breadth. See also Lerner (1994).

[21] According to the USPTO website: *"As of 12/31/2012, utility patent data, as distributed by year of application, are approximately 95% complete for utility patent applications filed in 2004, 89% complete for applications filed in 2005, 80% complete for applications filed in 2006, 67% complete for applications filed in 2007, 49% complete for applications filed in 2008, 36% complete for applications filed in 2009, and 19% complete for applications filed in 2010; data are essentially complete for applications filed prior to 2004."*

- A weighted count of patents based on *generality*. We base our definition of patent generality on the 4-digit International Patent Classification (IPC) following the definition in Hall et al. (2001). Generality of a patent is taken to be equal to one minus the Herfindahl index from all the technological classes that cite the patent. Formally, the generality index G_{it} of a patent i whose application date is t is equal to:

$$G_{it} = 1 - \sum_{j=1}^{J} \left(\frac{s_{j,t,t+5}}{\sum_{j=1}^{J} s_{j,t,t+5}} \right)^2 ,$$

  where $s_{j,t,t+5}$ is the number of citations received from other patents in ICP class $j \in \{1..J\}$ within five years after t. If the citing patent is associated with more than one technology class, we include all these classes to compute the generality index.

These measures have been aggregated at the state level by taking the sum of the quality measures over the total number of patents granted for a given state and a given application year and then divided by the population in the state. These different measures of innovation display consistent trends: hence the four states with the highest flows of patents between 1975 and 1990 are also the four states with the highest 5-year window citation counts, and similarly for the four most innovative states between 1990 and 2010 (California, New York, Massachusetts and Texas). From Figure 2, those states which experienced the fastest growth in innovation are Idaho, Washington, Oregon and Vermont; on the other hand, the states with the lowest growth in innovation are West Virginia, Oklahoma, Delaware and Arkansas. More statistics are given in Tables 1 and 2.

3.1.3 Control variables

When regressing top income shares on innovation, a few concerns may be raised. First, the state-specific business cycle is likely to have direct effects on innovation and on top income share. Second, top income share groups are likely to involve to a significant extent individuals employed by the financial sector (see for example Philippon and Reshef, 2012). In turn, the financial sector is sensitive to business cycles and it may also affect innovation directly. To address these two concerns, we control for the business cycle via the unemployment rate and for the share of GDP accounted for by the financial sector per inhabitant. In addition, we control for the size of the government sector which may also affect both top income inequality and innovation. To these we add usual controls, namely GDP per capita and the growth of total population. The corresponding data, namely on GDP, unemployment, total population and the share of the financial and public sectors, can be found in the Bureau of Economic Analysis (BEA) regional accounts.[22]

3.2 Estimation strategy

We seek to look at the effect of innovation measured by the flow of patents granted by the USPTO per inhabitants and by the quality of innovation on top income shares. We thus regress the top 1% income share on our measures of innovation. Our estimated equation is:

$$\log(y_{it}) = A + B_i + B_t + \beta_1 \log(innov_{i(t-2)}) + \beta_2 X_{it} + \varepsilon_{it}, \tag{15}$$

where y_{it} is the measure of inequality (which enters in log), B_i is a state fixed effect, B_t is a year fixed effect, $innov_{i(t-2)}$ is innovation in year $t-2$ (which enters in log as well),[23] and X is a vector of control variables. We discuss further dynamic aspects of our data later in the text. By including state and time fixed effects, we are eliminating permanent cross state differences in inequality and also aggregate changes in inequality.[24] We are essentially studying the relationship between the differential growth in innovation across states with the differential growth in inequality. In addition, by taking the log in both innovation and inequality, the coefficient β_1 can then be seen as the elasticity of inequality with respect to innovation.

Since we are using two-year lagged innovation on the right-hand side of the regression equation, and given what we said previously regarding the truncation bias towards the end of the sample period, we were able to run the regressions corresponding to equation (15) for t between 1977 and 2011 when measuring innovation by the number of patents and from 1977 and 2008 when measuring innovation using the quality-adjusted measures.

In all our regressions, we compute autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. By examining the estimated residual autocorrelations for each of the states we find that there is no significant autocorrelation after two lags. For this reason we choose a bandwidth equal to 2 years in the Newey-West

[22] Data description is given in Table 3.

[23] When *innov* is equal to 0, computing log(*innov*) would result in removing the observation from the panel. In such cases, we proceed as in Blundell *et al.* (1995) and replace log(*innov*) by 0 and add a dummy equal to one if *innov* is equal to 0. This dummy is not reported.

[24] We note that, after removing state and time effects, the inequality and innovation series are both stationary. For example, when we regress the log of the top 1% income share on its lagged value we find a precisely estimated coefficient of .821. Similarly when we regress innovation measured by citations in a 5-year window, on its one year lagged value, we find a precisely estimated coefficient of .779.

standard errors.[25]

4 Results from OLS regressions

In this section we present the results from OLS regressions of top income and other measures of inequality on innovation. We first look at the correlation between innovation and top income inequality. Then we look at the correlations between top income and other measures of inequality. Next, we look at how top income inequality correlates with innovation at different lags. Then we look at how the correlation between innovation and top income inequality is affected by the intensity of lobbying, and finally we look at the relationship between innovation and entrant versus incumbent innovation.

4.1 Innovation and top income inequality

Table 4 regresses the top 1% income share on our measures of innovation. The relevant variables are defined in Table 3. Column 1 uses the number of patents as a measure of innovation, column 2 uses the number of citations in a 5 year window, column 3 uses the number of claims, column 4 uses the generality weighted patent count and columns 5 and 6 use the number of patents among the top 5% and top 1% most cited patents in the year. All these values are divided by the population in the state, taken in log and lagged by 2 years.

From Table 4 we see that the coefficient of innovation is always positive and significant at the cross state level except when we use the number of patents per capita (column 1). This in turn suggests that particularly the more highly cited patents are associated with the top 1%, as those are more likely to protect true innovations. This is in line with Hall *et al.* (2005) who show that an extra citation increases the market share of the firm which owns the patent. Finally, the positive coefficient on the relative size of the financial sector reflects the fact that the top 1% involves a disproportionate share of the population working in that sector.

Because our measures or innovation and inequality are both taken in log, we can interpret the coefficient on innovation as an elasticity: namely, a 1% increase in the number of citations per capita is associated with a 0.3% increase of the top 1% income share. Moreover, we can compare the magnitude of this correlation with the correlation between the top 1% income share and the importance of the financial sector: thus a one standard deviation increase in

[25]The limited residual autocorrelation and the length of the time series (T is roughly equal to 30) justifies the use of a Newey-West estimator but we also present the main OLS regressions with clustered standard errors in Table B2 in Appendix B.

our measure of innovation leads to a 0.037 point increase in the log of top 1% income share whereas a one standard deviation increase in the share of financial sector in total GDP is associated with a 0.020 point increase in the log of top 1% income share.

4.2 Innovation and other measures of inequality

We now perform the same regressions as before but using broader measures of inequality: the top 10% income share, the Gini coefficient, the Atkinson index and the Theil index which are drawn from Frank (2009). Moreover, with data on the top 1% income share, we derive an estimate for the Gini coefficient of the remaining 99% of the income distribution, which we denote by $G99$ where:

$$G99 = \frac{G - top1}{1 - top1},$$

where G is the global Gini and top1 is the top 1% income share. In order to check if the effect of innovation on inequality is indeed concentrated on the top 1% income, we compute the average share of income received by each percentile of the income distribution from top 10% to top 2% and compare the coefficient on the regression of innovation on this variable with the one obtained with the top 1% income share as left hand side variable. This average size is equal to:

$$Avgtop = \frac{top10 - top1}{9},$$

where top10 represents the size of the top 10% income share.

Table 5 shows the results obtained when regressing these other measures of inequalities on innovation quality. We chose to present results for the citation variable but results are similar when using other measures of innovation quality. Column 1 reproduces the results for the top 1% income share. Column 2 uses the *Avgtop* measure, column 3 uses the top 10% income share, column 4 uses the overall Gini coefficient and column 5 uses the Gini coefficient for the bottom 99% of the income distribution to measure income inequality on the left-hand side of the regression equation. Column 6 uses the Atkinson Index with parameter 0.5.

We see from Table 5 that innovation: (a) is most significantly correlated with the top 1% income share; (b) is less (but still) correlated with the top 10% income share or with the average share between 10% and 1%; (c) is not significantly correlated with the Gini index and is negatively correlated with the bottom 99% Gini (although this negative effect is small).[26] Finally, the Atkinson index with coefficient equal 0.5 is positively correlated with

[26] This in turn may partly reflect the fact that, by concentrating market power within a few firms, innovation reallocates some rents from relatively high-earners towards very high-earners. For instance, in the context of our model, one could imagine that in the absence of innovation, a few firms behave as an oligopoly charging

innovation.

Finally, using new data recently released by Frank (2009), we were able to look at the effect of innovation on the very top of the income distribution, namely the top 0.01, 0.05 and 0.1% income shares. The correlation between innovation and top income share increased as we move to up the income distribution, with the coefficient of innovation reaching 0.065 for the top 0.01% income share. These results are presented in Table B3 of Appendix B.

4.3 Top income inequality and innovation at different time lags

One may first question the choice of two-year lag innovation in our baseline regression equation. In fact, two years is roughly the average time between a patent application and the date at which the patent is granted. For example, using Finnish individual data on patenting and wage income, Toivanen and Vaananen (2012) find an average lag of two years between patent application and patent grant, and they find an immediate jump in inventors' wages after patent grant. Other empirical results in two recent papers by Depalo and Di Addario (2014) and Bell *et al.* (2015) support the view that income can even peak before the patent is granted: Depalo and Di Addario (2014) find that inventors' wage peak around the time of the patent application, and Bell *et al.* (2015) show that the earnings of inventors start increasing before the filing date of the patent application. More generally, patent applications are mostly organized and supervised by firms who start paying for the financing and management of the innovation right after (or even before) the application date as they anticipate the future profits from the patent. Also, firms may sell a product embedding an innovation before the patent has been granted, thereby already appropriating some of the profits from the innovation.

Table 6 shows results from regressing top income inequality on innovation at various lags. We let the time lag between the dependent variable and our measure of innovation vary from 1 to 6 years. In order to have comparable estimates based on a similar number of observations, we chose to restrict the time period to 1981-2008. From this table, we see that the effect of lagged innovation is significant up to three-years lags, but with more lags, the effect becomes insignificant. This latter finding is consistent with the view that innovation should have a temporary effect on top income inequality due to imitation and creative destruction, in line with the Schumpeterian model in Section 2. Finally, the positive coefficient on one-year lagged innovation is in line with Depalo and Di Addario (2014) and

the mark-up η_L and dividing the profits among themselves. The owners of these firms would be high income earners but not necessarily in the top 1%. When innovation occurs, the leader captures all the rents and reaches the top 1% while the other individuals return to the production sector and see their income decline.

Bell *et al.* (2015) who argue that the effect of innovation on income should peak around the year of application.

4.4 Lobbying as a dampening factor

To the extent that lobbying activities help incumbents prevent or delay new entry, our conjecture is that places with higher lobbying intensity should also be places where innovation has lower effects on the top income share and on social mobility.

Measuring lobbying expenditures at the state level is not straightforward, in particular because lobbying activities often occur nationwide. To obtain a local measure of lobbying we use national sectoral variations in lobbying together with local variations in the sectoral composition in each state. More specifically, the OpenSecrets project[27] provides sector-specific lobbying expenditures at the national level for the period 1998-2011. To measure lobbying intensity at the state level, we construct for each state a Bartik variable, as the weighted average of lobbying expenditures in the different sectors (2-digits NAICS sectors), with weights corresponding to sector shares in the state's total employment from the US Census Bureau.

More precisely, we want to compute $Lob(i,.,t)$ the lobbying expenditure in state i in year t, knowing only the national lobbying expenditure $Lob(.,k,t)$ by sector k. We then define the lobbying intensity by sector k in state i at year t as:

$$Lob(i,k,t) = \frac{emp(i,k,t)}{\sum_{j=1}^{I} emp(j,k,t)} Lob(.,k,t),$$

where $emp(i,k)$ denotes industry k's share of employment in state i (where $1 \leq k \leq K$ and $1 \leq i \leq I$). From this we compute the aggregate lobbying intensity in state i as:

$$Lob(i,.,t) = \frac{\sum_{k=1}^{K} emp(i,k,t) Lob(i,k,t)}{\sum_{k=1}^{K} emp(i,k,t)}.$$

We then compute our measure of lobbying intensity by dividing the above measure of aggregate lobbying by the state population at year t. Table 7 shows the results from the OLS regression of the top 1% income share on innovation, our measure of lobbying intensity

[27] Data can be found in the OpenSecrets website

and the interaction between the two. Due to the limited time range for the lobbying data, we were able to run the regression only for the period 1998-2008. The results show that the overall effect of innovation on the top 1% income share is always positive and significant, the effect is weaker and even negative in states with higher lobbying intensity.

4.5 Entrants and Incumbents Innovation

Our empirical results so far have highlighted the positive relationship between innovation and top income inequality. In order to distinguish between incumbent and entrant innovation in our data, we rely on the work of Lai *et al.* (2013) which allows us to track the inventor(s) or assignee(s) for each patent over the period 1975-2010. We declare a patent to be an "entrant patent" if the time lag between its application date and the first patent application date of the same assignee amounts to less than 3 years.[28] We then aggregate the number of "entrant patents" as well as the number of "incumbent patents" at the state level from 1980 to 2010.[29]

According to our definition of an "entrant" innovation, 17% of patent applications from 1980 to 2010 correspond to an "entrant" innovation (this number increases up to 23.7% when we use the 5-year lag threshold to define entrant versus incumbent innovation). These "entrant" patents have more citations than the "incumbent" patents: for example in 1980, each entrant patent has 11.4 citations on average whereas an incumbent patent only has 9.5 citations, confirming the intuitive idea that entrant patents correspond to more radical innovations (see Akcigit and Kerr, 2010).

Table 8 presents the results from the regression of the top 1% income share over incumbent and entrant innovation, where these are respectively measured by the number of patents per capita in columns 1, 2 and 3 and by the number of citations per capita in columns 4 to 6. The coefficients on entrant innovation are always positive and significant, and in the horse race regressions of top inequality on incumbent and entrant innovation (columns 3 and 6), only the coefficients for entrant innovation come out significant although the difference between the coefficients for entrant and incumbent innovation are not statistically significant.[30]

[28] We checked the robustness of our results to using a 5-year lag instead of a 3-year lag threshold to define entrant versus incumbent innovation (see Table B4). Here we only focus on patents issued by firms and we have removed patents from public research institutes or independent inventors.

[29] We start in 1980 to reduce the risk of wrongly considering a patent to be an "entrant" patent just because of the truncation issue at the beginning of the time period. In addition, we consider every patent from the USPTO database, including those with application year before 1975 (but which were granted after 1975).

[30] Because the data of Lai *et al.* (2013) stops in 2010, we limit the sample period for the panel regressions to 1980-2004.

4.6 Summary

The OLS regressions of innovation on income inequality performed in this section lead to interesting correlation results that are broadly in line with the Schumpeterian view developed in the model, namely: (i) innovation is positively correlated with top income inequality; (ii) innovation is not significantly correlated with broader measures of inequality (Gini,...); (iii) the correlation between innovation and top income inequality is temporary (lagged innovation ceases to be significant when the lag becomes sufficiently large); (iv) the correlation between innovation and top income inequality is lower in states with higher lobbying intensity (v) top income inequality is positively correlated with both, entrant and incumbent innovation.

5 Endogeneity of Innovation and IV Results

In this section we argue that the positive correlations between innovation and top income inequality uncovered in the previous section, at least partly reflect a causal effect of innovation on top income. To reach this conclusion we have to account for the possible endogeneity of our innovation measure. Endogeneity could occur through the feedback of inequality to innovation. For example, a growth in top incomes may allow incumbents to erect barriers against new entrants thereby reducing innovation and inducing a downward bias on the OLS estimate of the innovation coefficient. We develop this point further below.

Our first instrument for innovation exploits changes in the state composition of the Appropriation Committee of the Senate which allocates federal funds in particular to research across US states. Then, we show that this Appropriation Committee instrument can be combined with a second instrument which explores knowledge spillovers across states.

5.1 Instrumentation using the state composition of appropriation committees

We instrument for innovation using information on the time-varying state composition of the appropriation committee. To construct this instrument, we gather data on membership of these committees over the period 1969-2010 (corresponding to Congress numbers 91 to 111).[31] The rationale for using this instrument is that the appropriation committee allocates federal funds to research education across US states. Even though the appropriation committee

[31] Data have been collected and compared from various documents published by the House of Representative and the Senate. The name of each congressman has been compared with official biographical informations to determine the appointment date and the termination date.

is not explicitly dedicated to research and research education, an important fraction of the federal funds it allocates across states goes to research education. A member of Congress who sits in such a Committee often pushes for earmarked grants aimed at subsidizing research education in the state in which she has been elected, in order to increase her chances of reelection in that state. Consequently, a state with one of its congressmen seating on the committee is likely to receive more funding and to develop its research education, which should subsequently increase its innovation in the following years.

Aghion et al (2009) note that "research universities are important channels for pay back because they are geographically specific to a legislator's constituency. Other potential channels include funding for a particular highway, bridge, or similar infrastructure project located in the constituency". Moreover, in Table 8 of their paper, they show that among all categories of non-education federal expenditures, only expenditures on highways show a positive correlation with education federal expenditures. In addition, the OpenSecrets project website lists the main recipients of the 111$^{\text{th}}$ Congress Earmarks in the US (between 2009 and 2011), and universities rank at the top together with defense companies. We shall control for state-level highway and military expenditures in our IV regressions as detailed below.

Changes in the state composition of the Appropriation Committee have little to do with growth or innovation performance in those states. Instead, they are determined by events such as anticipated elections or more unexpectedly the death or retirement of current heads or other members of these committees, followed by a complicated political process to find suitable candidates. This process in turn gives large weight to seniority considerations with also a concern for maintaining a fair political and geographical distribution of seats. In addition, legislators are unable to fully evaluate the potential of a research project and are more likely to allocate grants on the basis of political interests. Both explain why it is reasonable to see the arrival of a congressman in the appropriation committee, as an exogenous shock on innovation (a decrease in θ_E and θ_I in the context of our model).

Based on these Appropriation Committee data, different instruments for innovation can be constructed. We follow the simplest approach which is to take the number of senators (0, 1 or 2) or representatives who seat on the committee for each state and at each date.

A related concern is that the composition of the appropriation committee would reflect the disproportionate attractiveness of states such as California and Massachusetts. However, other states have been well represented on the committee -for example Alabama had one senator, Richard C. Shelby, sitting on the Committee between 1995 and 2008-, whereas California had no committee members until the early 1990s. [32] Also, if we look at the cross-

[32] More statistics on the state composition of the Senate Appropriation Committee is provided in Table 9.

state allocation of earmarks from the 111$^{\text{th}}$ Congress as shown on the OpenSecrets website, we see that the states that received the highest amount of earmarks per inhabitant, are Hawai (not too surprising, since the Chairman of the Senate Appropriation Committee at the time, Daniel K. Inouye was himself a senator from Hawaii) and North Dakota. Finally, any given state cannot have more than two representatives on the Senate committee.

Next, we need to find the appropriate time-lag between a congressman's accession into the appropriation committee and the effect this may have on innovation. We chose to instrument innovation by committee composition with a lag of two or three years, which adds to the two-year lag between innovation and top income inequality in the baseline regression.[33]

Although changes in the composition of the Appropriation Committee can be seen as exogenous shocks to innovation across states, there is still a concern about potential direct effects of such changes on the top 1% income share that do not relate to innovation. There is not much data on appropriation committee earmarks; yet, for the years 2008 to 2010, the Taxpayers for Common Sense, a nonpartisan budget watchdog, reports data on earmarks in which we can see that infrastructure, research, education and military are the three main recipients for appropriation committees' funds. In addition, when looking more closely at top recipients, we find that most are either universities or defense-related companies.[34] One can of course imagine a situation in which the (rich) owner of a construction or military company will capture part of these funds. In that case, the number of congressmen seating in the committee of appropriation would be correlated with the top 1% income share, but for reasons having little to do with innovation. To deal with such possibility, we use data on total federal allocation to states by identifying the sources of state revenues. Such data can be found at the Census Bureau on a yearly basis. Using this source, we identify for each state, military expenditures and a particular type of infrastructure spending, namely highways, for which we have consistent data from 1975 onward. We control for both in our regressions.

Table 10 shows the results from the IV regression of top income inequality on innovation, using the state composition of the Senate appropriation committee as the instrumental vari-

[33] Yet, one may wonder how changes in the Appropriation Committee of the Senate could affect top income inequality in the states already after four or five years. First, as pointed out by Aghion *et al.* (2009), research education funding in a state is immediately affected when representation of that state in the Appropriation Committee changes. Second, research grants often reward research projects that are already completed. Third, changes in research grants induce quick multiplier effects in the private sector (this is in line with Toole, 2007, who shows that in the pharmaceutical industry, the positive impact of public R&D on private R&D is the strongest after 1 year).

[34] Such data can be found on the Opensecrets website

able for innovation.[35,36] Column 1 uses the number of patents as a measure of innovation, column 2 uses the number of citations in a 5 year window, column 3 uses the number of claims, column 4 uses the generality weighted patent count and columns 5 and 6 use the number of patents among the top 5% and top 1% most cited patents in the year. In all cases, the instrument is lagged by 3 years with respect to the innovation variable it is instrumenting (and recall that innovation is itself lagged by 2 years in the main regression). In all cases, the resulting coefficient on innovation is positive and significant. Moreover, with the exception of columns 4 and 6, the F-statistics is above 10 suggesting that our instrument is reasonably strong.

Now, regarding the magnitude of the impact of innovation on top income inequality implied by Table 10, we see that an increase of 1% in the number of patents per capita increases the top 1% income share by 0.24% (see column 1 in Table 10) and that the effects of a 1% increase in the citation-based measures are of comparable magnitude. This means for example that in California where the flow of patents per capita has been multiplied by 3.1 and the top 1% income share has been multiplied by 2.4 from 1980 to 2005, the increase in innovation can explain 30% of the increase in the top 1% income share over that period. On average across US states, the increase in innovation as measured by the number of patents per capita explains about 24% of the total increase in the top 1% income share over the period between 1980 and 2005. Looking now at cross state differences in a given year, we can compare the effect of innovation with that of other significant variables such as the importance of the financial sector. Our IV regression suggests that if a state were to move from the first quartile in terms of the number of patents per capita in 2005 to the fourth quartile, its top 1% income share would increase on average by 3.5 percentage points. Similarly, moving from the first to the fourth quartile in terms of the number of citations, increases the top 1% income share by 3.3 percentage points. By comparison, moving from the first quartile in terms of the size of the financial sector to the fourth quartile, would lead to a 4.5-percentage-point increase in the top 1% income share.[37]

[35]The results from the first stage regression and the reduced form regression, are shown in Table B5 in Appendix B.

[36]As we have a long time series for each state, we are not concerned about 'short T' bias in panel data IV. We apply instrumental variables estimator directly to time and fixed effects regression equation (15).

[37]Yet, one should remain cautious when using our regressions to assess the true magnitude of the impact of innovation on top income inequality, as there are reasons to believe our regression coefficients may either overestimate or underestimate that impact. Underestimate: (i) the number of citations has increased by more than the number of patents over the past period, which suggests that the effect of innovation on top income inequality is greater than 24%; (ii) if successful, an innovator from a relatively poor state, is likely to move to a richer state, and therefore to not contribute to the top 1% share of her own state; (iii) an innovating firm may have some of its owners and top employees located in a state different from that of inventors, in which case the effect of innovation on top income inequality will not be fully internalized by

5.2 Discussion

The following concerns could be raised by this regression. First, it could be that some of our control variables are endogenous and that, conditional upon them, our instruments may be correlated with the unobservables in our model. To check that our results are robust to this possibility, we re-run our IV regressions, with state and year fixed effects but removing the control variables. And in each case we find that the regression coefficients on the various measures of innovation remain of the same order of magnitude and significance compared to the corresponding IV regressions with all the control variables, but the corresponding first stage F-statistics are lower (between 7 and 9.3).[38]

Second, the magnitude of the innovation coefficients in the IV regression is larger than in the OLS regressions. One potential reason has to do with the relationship between innovation and competition. More specifically, suppose that the relationship between competition and innovation lies on the upward part of the inverted-U relationship between these two variables (see Aghion *et al.* 2005), and consider a shock to the level of competition faced by a leading firm, which increases its market power—such a shock could for example result from an increase in lobbying or from special access to a new enlarged market. This shock will increase the firm's rents which in turn should contribute to increasing inequality at the top. However, on this side of the inverted-U, this will also decrease innovation. Therefore, it induces an increase in top inequality that is bad for innovation. As it turns out, lobbying is indeed positively correlated with the top 1% income share and negatively correlated with the flow of patents. Relatedly, our model shows that a higher level of mark-ups for an incumbent who has failed to innovate can also lead to higher top income inequality and lower innovation; this higher mark-up level may in turn reflect slow diffusion of new technologies and/or high entry barriers.

Third, one might raise the possibility that some talented and rich inventors decide to move to states that are more innovative or to benefit from lower taxes. This would enhance the positive correlation between top income inequality and innovation although not for the

the state where the patent is registered. Overestimate: not all innovations are patented; if the share of innovations that get patented is increasing over time, then the increase in innovation will be less than the measured increase in patenting, so that we might in fact explain a little less than 22% of the increase in top 1% income share. Importantly, as long as the increase in the share of patented innovations is the same across states, this would not bias our regression coefficients (as this effect would be absorbed in the time fixed effect). Furthermore, Kortum and Lerner (1999) argue that the sharp increase in the number of patents in the 90's reflected a genuine increase in innovation and a shift towards more applied research instead of regulatory changes that would have made patenting easier.

[38]The key assumption here is that the unobservables in the model are mean independent of the instruments conditional on the included controls.

reason captured by our IV strategy.[39] However, building on Lai *et al.* (2013), we are able to identify the location of successive patents by a same inventor. This in turn allows us to delete patent observations pertaining to inventors whose previous patent was not registered in the same state. Our results still hold when we look at the effect of patents per capita on the top 1%, with a regression coefficient which is essentially the same as before.

5.3 Other IV results

In Appendix B we show the results from replicating in IV the OLS regressions of Section 4. First, regressing broader measures of inequality on innovation, we find that innovation has a positive impact on top income shares but not on Gini coefficients (Table B6). Note that the effect of innovation on the top 10% remains positive but is no longer significant. Second, regressing top income inequality on innovation at various lags, we find that the effect of lagged innovation is strongest after 2 years, although it is already significant after 1 year; after 4 years or more, the effect becomes smaller and insignificant (Table B7). These latter findings confirm those in the corresponding OLS Table 6, and speak again to the fact that innovation has a temporary effect on top income inequality.

6 Robustness checks

In this section we discuss the robustness of our basic regression results to introducing a second instrument which exploits knowledge spillovers across states, and to adding more controls. Table 11 shows the results from the IV regression where we combine the appropriation committee and the spillover instruments. Table 12 shows the results from adding various controls to the OLS regressions.

6.1 Adding a second instrument

To add power to our instrumental variable estimation, here we combine it with a second instrument which exploits knowledge spillovers across states. The idea is to instrument innovation in a state by its predicted value based on past innovation intensities in other states and on the propensity to cite patents from these other states at different time lag. Citations reflect past knowledge spillovers (Caballero and Jaffe 1993), hence a citation network reflects

[39] Moretti and Wilson (2014) indeed show that in the biotech industry, the decline in the user cost of capital in some US states induced by federal subsidies to those states, generated a migration of star scientists into these states.

channels whereby future knowledge spillovers occur. Knowledge spillovers in turn lower the costs of innovation (in the model this corresponds to a decrease in θ_I or θ_E). To build this predicted measure of innovation, we rely on the work of Acemoglu *et al.* (2016) and integrate the idea that the spillover network can be very different when looking at different lags between citing and cited patent. We thus compute a matrix of weights $w_{i,j,k}$ where for each pair of states (i, j) and for each lag k between citing and cited patents where k lies between 3 and 10 years,[40] $w_{i,j,k}$ denotes the relative weight of state j in the citations with lag k of patents issued in state i, aggregated over the period from 1975 to 1978. [41]

Using this matrix, we compute our instrument as follows: if $m(i, j, t, k)$ is the number of citations from a patent in state i, with an application date t to a patent of state j filed k years before t, and if $innov(j, t-k)$ denotes our measure of innovation in state j at time $t-k$, then we posit:

$$w_{i,j,k} = \frac{\sum_{t=1975}^{1978} m(i,j,t,k)}{\sum_{t=1975}^{1978} \sum_{l \neq i} m(i,l,t,k)} \; ; \; KS_{i,t} = \frac{1}{Pop_{-i,t}} \sum_{k=3}^{10} \sum_{j \neq i} w_{i,j,k} innov(j, t-k),$$

where $Pop_{-i,t}$ is the population of states other than state i and the log of KS is the instrument. To reduce the risk of simultaneity, we set a one year time lag between the endogenous variable and this instrument. Without normalizing by $Pop_{-i,t}$, our measure of spillovers would mechanically put at a relative disadvantage a state which is growing relatively faster than other states.

Reverse causality from top income inequality to this knowledge spillover IV seems unlikely (the top 1% income share in one state is unlikely to cause innovations in other states).[42] However, one may worry that this instrument would capture regional or industry trends that are not directly the result of innovation and yet affect both top income inequality and innovation in that state. Consider for example two states that are highly intensive in terms of, say, the computer sector, then a demand shock in this sector may boost innovation and the top 1% income share, violating our exclusion restriction. To capture such demand shocks, we

[40] Looking at all the patents in our sample granted over the period from 1975 to 2014, we find that 67% of backward citations are made to patents filed less than 10 years before the citing patent.

[41] We observe all the patents which received citations from patents granted after 1975 even if the cited patents were granted before 1975.

[42] Yet, reverse causality might arise from the same firm citing itself across different states. We check that this has, if anything, a very marginal effect by removing citations from a firm to itself in two different states when constructing the weights: the results are essentially unaffected by this change.

add two additional control variables in our regressions. First, using an average of the weights calculated before for k between 3 and 10, we calculate a weighted average of other states' per capita GDP. Second, we build new weights based on the angular distance between two states' industry compositions in the manufacturing sector. These new weights are averaged over a three-year window. Using these industry-composition-based weights, we compute a weighted sum of innovation in other states and divided this sum by $Pop_{-i,t}$. Second, by looking at past innovation that occurred at least 3 year earlier, we set a sufficiently long time lag to reduce the risk of capturing demand shocks that are much faster.

Finally, an overidentification test that uses the spillover and appropriation committee instruments, does not reject the validity of the instruments: indeed, the p-value associated with the null hypothesis is always larger than 10% (note however that when Claims and Generality measures are used, this p-value is a little below 10%). This in turn reinforces the first instrument.[43]

Table 11 presents the results from the IV regressions of top income inequality on the two instruments combined.[44][45] As in Table 10, the coefficients are always positive and significant (now at the 1% level). This is all the more remarkable that the two instruments are uncorrelated once one controls for states and time fixed effects. The F-statistics for the two instruments combined, are always above 10.

6.2 The role of finance and natural resources

When considering top income shares and other inequality measures on the one hand and innovation on the other hand, we abstracted from industry composition in the various states. However, two particular sectors deserve to be considered more closely: Finance and Natural resources.

The financial sector is overrepresented in the top 1% income share (even though most individuals in the top 1% do not work in the financial sector). More specifically, Guvenen *et al.* (2014) find that 18.2% of individuals in the top 1% work in the Finance, Insurance and Real Estate sector (versus 5.3% for the rest of the population), and that these individuals' income is particularly volatile. To make sure that our effects are not mainly driven by the financial sector, in the above regressions we already controlled for the share of the financial

[43]This also deals with the potential objection that innovation in other states $j \neq i$ could have a direct impact on productivity in state i, and thereby directly affect top incomes in that state. If that were the case, the two instruments combined would be correlated with the error term and therefore in that case the overidentification test would reject the null hypothesis.

[44]In the Appendix, Table B8, we show the results from the IV regressions using only the second instrument.

[45]The results from the corresponding first stage and reduced form regressions, are shown in Table B5 in the Appendix B.

sector in state GDP. Here, we perform additional tests. First, we add the average employee compensation in the financial sector as a control to capture any direct effect an increase in financial sector's employee compensation might have on the top 1% income share (see column 1 of Table 12). Second, we exclude states in which financial activities account for a large fraction of GDP. We selected four such states: New York, Connecticut, Delaware and Massachusetts (see column 2 of Table 12). Third, financial innovations themselves might directly increase rents and therefore the top 1% income share. To account for this latter channel, we subtract patents belonging to the class 705: "Financial, Business Practice" related to financial activities in order to exclude innovations in the financial sector (see column 3 of Table 12). The regressions of the top 1% income share on innovation corresponding to these three robustness tests uses the number of citations per capita within a 5-year window to measure of innovation. In each case, the effect of innovation on the top 1% income share is significant and positive, showing very stable values when moving from one specification to another. Another potential issue related to finance is that financial development should impact both innovation (by providing easier access to credit to potential innovators) and income inequality at the top (by boosting high wages). Here we construct a variable specifically designed to directly capture this channel. For each US state, we divide the number of patent applications in that state into 16 NAICS categories mapped with patent technological classes from the USPTO[46] and we use the external financial dependence numbers computed by Kneer (2013) and averaged over the period 1980-1989. External financial dependence is defined as the ratio of capital expenditure minus cash flow divided by capital expenditure (see Rajan and Zingales, 1998). We multiply the number of patents in each NAICS sector in that state by that index and then divide by the total number of patents to compute a variable representing the level of financial dependence of innovation for each state. This variable (denoted EFD in Table 12) should capture a variation in innovation at state-level driven by a sector that is highly dependent on external finance. Results for regressing the top 1% income share on the number of citations per capita within a 5 year window when controlling for EFD are presented in column 4. We see that the effect of innovation remains significant and the coefficient is slightly lower than the corresponding coefficient when we do not control for EFD. In addition, the effect of the financial dependence of innovations on the top 1% income share is positive and significant.

Natural resources and particularly oil extraction represent a large share of GDP in certain states (In Wyoming, West Virginia and particularly Alaska, oil extraction activities account for almost 30% of total GDP in 2010), so that in these states the top 1% income share is

[46]The latest version of this mapping can be found in the USPTO website.

likely to be affected by these sectors which are quite volatile (oil extraction is highly sensitive to energy prices fluctuation). To deal with this concern, we control for the share of natural resources in GDP. In addition, we first add the share of oil extraction related activities in state GDP as a control variable; and second, we remove patents from class 208 (Mineral oils: process and production) and 196 (Mineral oils: Apparatus). Results are presented in columns 5 and 6 of Table 12. Here again, our results remain significant.

Finally Table B9 shows similar results from performing the same robustness tests in IV regressions with the appropriation committee instrument (as with the other robustness checks below, the results are similar when we use the other instrument or both).

6.3 Accounting for changes in top tax rates

Taxation is likely to affect both innovation incentives and the 1% income share. In particular, high top marginal income tax rates may reduce efforts by top earners, divert their pay from wages to perks, and reduce their incentives to bargain for higher wages (see, in particular, Piketty *et al.*, 2014). In this subsection, we address this concern more directly.

More specifically, we use data from the NBER TAXSIM project. This database provides information on marginal tax rates for various levels of incomes ($10000, $25000, $50000, $75000 and $100000 yearly incomes) and for labor, capital and interest incomes from 1977 onward. We use the state marginal labor income tax rate for individuals earning $100000 per year as an additional control when regressing the top 1% income share on innovation. The results are displayed in column 7 of Tables 12 and B9: the effect of innovation on the top 1% income share remains positive and significant.

6.4 Looking at industry composition

We now check that our results are robust to controlling for the sectoral composition of innovation. First, we use the mapping between patent technological classes and NAICS sectors to remove patents related to category 334: "Computer and Electronic Products", to deal with the concern that the effect of innovation on top income inequality might be concentrated in the fast-growing computer industries. Similarly, we remove patents from the pharmaceutical sector (NAICS 3254) and from the electrical equipment sector (NAICS 335). In each case, we conduct our preferred regression using the number of citations within a five-year window to measure innovation. The coefficient on innovation remains quite stable across all these specifications. Next, in our regressions we add controls for the share of three

NAICS sectors.[47] Innovation remains positively and significantly correlated with the top 1% income share in all our regressions.

Then, we use the COMTRADE database to look at the extent to which our effect of innovation on top income inequality is driven more by more export-intensive sectors. Over the period from 1975 to 2013, we identified three sectors that are particularly export-intensive: Transportation, Machinery and Electrical Machinery. When we regress the top 1% income share on the number of citations within a five-year window without counting citations to patents in these sectors, we find a positive and significant coefficient. All these results are shown in the Appendix B, Table B10 in OLS and Table B11 in IV.

6.5 Controlling for agglomeration effects

One may wonder whether our results do not reflect potential agglomeration effects: for example, suppose that some exogenous investment taking place in one particular location (think of the Silicon Valley), makes that location become more attractive to skilled/talented individuals from other parts of the US. Then the resulting increased agglomeration of high-skill individuals in that location, should result in both, a higher share in the top 1% income share and in an increase in innovation in the corresponding US state, but without the former necessarily resulting from the latter.

Looking at Figure 2 in the introduction hints at the fact that this should not be such a big concern: in particular we see that neither California nor Massachusetts are among the states that show the fastest increase in both, innovation and top income inequality, over the period we analyze.

To address the agglomeration objection head on, we proceed as follows: in any state i at any date t, we look at the most, the two most and the three currently most innovative technological classes from our patent dataset in that state. We then compute the number of patents in these technology classes in that state in that year. The log of that number is our new control variable $Agglo_{it}$ which is meant to capture potential agglomeration effects in state i in year t.

Running our previous regressions with these additional control variables turns out not to affect our results as seen in Table B12 in the Appendix B (see the first three columns of tables for OLS and the three other columns for IV regression results).

[47] In order to obtain complete series, we replace the pharmaceutical sector by the whole chemistry manufacturing sector (NAICS 325).

7 Innovation and social mobility

In this section we consider the relationship between innovation and social mobility

7.1 From cross-state to CZ-level analysis

Panel data on social mobility in the United States are not (yet) available. Therefore, to study the impact of innovation on social mobility without reducing the number of observations too much, we move from cross-state to cross-commuting zones (CZ) analysis and use the measures of social mobility from Chetty *et al.* (2014). A commuting zone (CZ) is a group of neighboring counties that share the same commuting pattern. There are 741 commuting zones which cover the whole territory of the United States. Some CZs are in rural areas whereas others are in urban areas (large cities and their surroundings). At the CZ level, we do not have data on top income shares for the whole population. However, Chetty *et al.* (2014) use the 2000 census to provide estimates for the top 1% share as well as for the Gini index for a sample of adults at CZ and MSA level. Using that information, we compute cross-sectional measures of inequality as an average between 1996 and 2000. If we look at urban CZs, the three largest top 1% income shares are in New York (23.6%), San Jose (26.4%) and San Francisco (29.1%), all of which are highly innovative areas.

To associate a patent to a CZ location, we rely on Lai *et al.* (2013) to complete the USPTO database with assignee and inventor names. This enables us to associate each inventor with her address and her zipcode which can be linked up to a county, and ultimately to a commuting zone. Finally, we aggregate county level data on GDP and population from the Bureau of Economics Analysis (BEA) to compute GDP per capita and population growth. All other data are taken from Chetty *et al.* (2014).

Using all these data, we can first check whether the effects of innovation on the top 1% income share and on the Gini index are consistent with our cross-states findings. Table 13 displays the results from the regression when the logarithm of the number of patents per capita is used as a measure of innovation. We add controls for GDP per capita, for the growth of total population and for the size of local government proxied by the logarithm of the local government's total expenditure per capita. In addition, we add the share of the manufacturing sector, the labor force participation rate taken in 1996-2000, college graduation rate and the local expenditures in public school per student during the same period. Standard errors are clustered by state to account for potential correlation across neighboring CZs. As seen from the first two columns of Table 13, the effect of innovation on the top 1% income share is positive and significant (column 1) and robust to the addition of a dummy equal to one if the CZ belongs to urban areas (column 2). When regressing innovation on inequality as measured by the Gini coefficient and on the Gini coefficient for the bottom 99% of the income distribution, the coefficients are much smaller or even negative (columns 3 to 6). All these observations are consistent with our core results at the cross state level.

7.2 The effect of innovation on social mobility

Having moved from cross-state to cross-CZ analysis allows us to look at how innovation affects social mobility, using the various measures of social mobility in Chetty *et al.* (2014) combined with our local measures of innovation and with the various controls mentioned above. There, absolute upward mobility is defined as the expected percentile or "rank" (from 0 to 100) for a child whose parents belonged to some P percentile of the income distribution. Percentiles are computed from the national income distribution. The ranks are computed over the period 2011-2012 when the child is aged around 30 whereas the percentile P of parents income is calculated over the period between 1996 and 2000 when the child was aged around 15. Once again, the intensity of innovation in each CZ is measured by the average number of patents per capita, but this time, we take the averages over the period 2006-2010

One potential concern with these data for our purpose, is that social mobility is based on the location of the parents not the children, and therefore the data do not account for children who move to and then innovate in a different location from that of their parents. However, if anything this should bias our results downwards: if many individuals migrate out of a specific CZ to innovate in San Francisco or New York, this CZ will exhibit high social mobility but low innovation.

We thus conduct the following regression:

$$log(Mob_k) = A + \beta_1 log(innov_k) + \beta_2 X_k + \varepsilon_k,$$

where Mob_k is our measure of upward social mobility, and $innov_k$ is our measure of innovation (the number of patents per capita) for CZ k. We cluster standard errors by state.

Table 14 presents our results for this cross-section OLS regression, where we add the same set of control variables as in the previous subsection. Columns 1 and 4 look at the effect of innovation on upward mobility as measured by the child expected percentile in the income distribution at 30 when parent income belongs to the 25th percentile. The effect of innovation is positive and significant. Columns 2, 3, 5 and 6 show the effects of innovation on the probability for a child to belong to the highest quintile in income distribution at age

30 when her parent belonged to a lower quintile. The lower the quintile to which parents belonged, the more positive and significant is the correlation between innovation and upward mobility. If we continue with quintiles 3 and 4, the effect of innovation on social mobility is still significant for quintile 3 (but only when college per capita and manufacturing share are not included) and negative and not significant for quintile 4. Not surprisingly, school expenditures, colleges per capita and participation rate also play a positive role in explaining upward social mobility, while the size of the manufacturing sector is negatively correlated. Finally, column 7 shows the overall effect of innovation on upward mobility measured by the probability to reach the highest quintile when parent belonged to any lower quintile. Here again, the correlation is positive.

One concern is worth mentioning here: in some CZs, the size of the top quintile is very small, reflecting the fact that it is almost impossible to reach this quintile while staying in this CZ. This case often occurs in rural areas: for example, in Greenville, a CZ in Mississippi, only 7.5% of children in 2011-2012 (when they are 30) belong to the highest quintile in the national income distribution. To address this concern, we conduct the same regressions as above but we remove CZs where the top quintile has a size below 10% and below 15% (this exclude respectively 7 and 100 CZs). All our results remain consistent with columns 1 to 6 of the previous regressions.[48] In fact, the results are even stronger, with the coefficient of innovation being now always significant at the 5% level.

All the results presented in this section are consistent with the prediction of our model that innovation increases mobility at the top. Yet, we should bear in mind that these are just cross-sectional OLS correlations, and this remark holds for all other CZ level regressions in this section.

7.3 Lobbying, entrant versus incumbent innovation, and social mobility

Our empirical results so far have highlighted the positive association between innovation on the one hand and social mobility on the other hand. Now, recall that our model suggests that the effect of innovation on social mobility should operate mainly through entrant innovation, and that entry barriers should dampen that effect.

To test these predictions, we conduct the same regression as in the previous section at

[48]This result is confirmed by performing the same regression on the whole sample of CZs but adding an interaction term between the number of patents per capita and a dummy equal to one if the CZ has a top quintile of size higher than 15% of total CZ population. The coefficient for this interaction term is positive and significant.

the cross CZ level but considering separately entrant innovation and incumbent innovation on the right hand side of the regression equation, where entrants and incumbents are defined as follows: an entrant patent (resp. an incumbent patent) is one where the assignee did not patent (resp. did patent) before 2006. Table 15 presents our results. Columns 1 to 3 regress our three measures of social mobility on the number of "entrant patents" per capita, whereas columns 4 to 6 regress the three measures of social mobility on the number of "incumbent patents". The positive and significant coefficients in the first three columns, as compared to columns 4 to 6, suggest that the positive effect of innovation on social mobility is mainly driven by new entrants. This conjecture is confirmed by the horse race regression in column 7 in which both entrant innovation and incumbent innovation are included as right-hand side variables. There, we clearly see that all the effect of innovation on social mobility is associated with entrant innovation.

We next look at how lobbying intensity interacts with the effect of innovation on social mobility using cross-MSA data. As explained above, we aggregated patent applications by zipcode and then by MSA and used mobility data from Chetty *et al.* (2014) who only provide absolute mobility data and no transition matrix for MSAs. Our regular control variables (GDP per capita, population growth, share of financial sector and government size) have been found in the BEA and averaged over the period 2006-2010. Overall, we are left with 352 MSAs which can be separated in two groups of equal size, respectively with high and low lobbying activities. Columns 1 and 2 of Table 16 show the effect of innovation as measured by the number of entrant patents per capita (in log) on the logarithm of absolute upward mobility. Column 1 focuses on MSAs above median in terms of lobbying activities and column 2 on other MSAs. Similarly, columns 3 and 4 look at the effect of the number of incumbent patents per capita on absolute upward mobility. We see that the effect of entrant innovation on social mobility is positive and significant only for MSAs that are below median in terms of lobbying intensity. In addition, incumbent innovation has no effect on social mobility, whether we look at MSAs above or below the median in terms of lobbying intensity. These results confirm the idea that lobbying dampens the impact of innovation on social mobility by reducing the effect of entrant innovation.

8 Conclusion

In this paper we have looked at the relationship between top income inequality and innovation. First, we found positive and significant correlations between measures of innovation on the one hand, and top income inequality on the other hand. We also showed that the correlations between innovation and broad measures of innovation are not significant, and that top income inequality is not correlated with highly lagged innovation. Second, we argued that these correlations at least partly reflect a causal effect from innovation to top income shares. Third, we showed that innovation is positively associated with social mobility.

Overall, our findings suggest interesting avenues for further research on (innovation-led) growth, inequality and social mobility.

A first extension would be to contrast innovation and other sources of top income inequality, for example from financial and lobbying activities, by looking at the effects of these other sources on other measures of inequality and on social mobility. Our conjecture is that, unlike innovation, lobbying should be positively correlated with broad measures of inequality, and negatively correlated with social mobility.

Relatedly, it would be interesting to link this work directly with the inverted-U results in Aghion *et al.* (2005), looking at how competition (or the lack of it) affects top income inequality both directly and through its effects on innovation, and examining the circumstances under which increasing top inequality may induce a reduction in competition and thereby depress innovative activity.

Another extension would be to explore policy implications. In particular, how do we factor in innovation in tax policy design, and how should we combine tax policy with other policy instruments (competition and entry policy, patent policy, R&D subsidies,...) to achieve more inclusive growth?

Another avenue for future research would be to look at the effect of innovation on top income inequality in cross-country panel data. Preliminary OLS regressions show a positive and significant correlation between our innovation measures and top 1% income share in cross-country panel.

A fifth extension would be to explore the relationship between innovation, top income inequality and social mobility using individual data on revenues and patenting.[49]

Finally, our results on the impact of lobbying suggest that the relationship between innovation and income inequality depends upon institutional factors which vary across countries. Further research should thus look deeper into how institutions affect the relationship between top income inequality and innovation. These and other extensions of the analysis in this paper are left for future research.

[49] In Aghion et al. (2015) such a study is conducted using Finnish individual data over the period 1990-2000. See also Toivanen and Vaananen (2012) and Bell *et al.* (2015).

References

Abrams, D., Akcigit, U. and Popadak, J. (2013), "Patent Value and Citations: Creative Destruction or Strategic Disruption?," *PIER Working Paper Archive 13-065*, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

Acemoglu, D (1998), "Why Do New Technologies Complement Skills: Directed Technical Change and Wage Inequality", *Quarterly Journal of Economics*, 113, 1055-1089

Acemoglu, D (2002), "Technical Change, Inequality, and the Labor Market", *Journal of Economic Literature*, 40, 7-72

Acemoglu, D (2007), "Equilibrium Bias of Technology", *Econometrica*, 75,1371-1409.

Acemoglu, D., and Robinson, J (2015), "The Rise and Decline of General Laws of Capitalism", *Journal of Economic Perspectives,* 29, 3-28.

Acemoglu, D., Akcigit, U. and Kerr, W (2016) "The Innovation Network," mimeo Harvard

Aghion, P., and Howitt, P (1992), "A Model of Growth Through Creative Destruction", *Econometrica*, 60, 323-351

Aghion, P., and Howitt, P (1997), *Endogenous Growth Theory,* MIT Press.

Aghion, P., and Tirole, J (1994), "On the Management of Innovation", *Quarterly Journal of Economics,* 109, 1185-1209.

Aghion, P., Bloom, N., Blundell, R., Griffith, R. and Howitt, P (2005), "Competition And Innovation: An Inverted-U Relationship," *Quarterly Journal of Economics*, 120, 701-728.

Aghion, P., Boustan, L., Hoxby, C., and Vandenbussche, J (2009), "The Causal Impact of Education on Economic Growth: Evidence from US", mimeo Harvard.

Aghion, P., Akcigit, U., Toivanen, O., and Vaananen, A (2015), "Living "American Dream" in Finland: The Social Mobility of Innovators", mimeo LSE

Akcigit, U., and Kerr, W (2010), "Growth Through Heterogeneous Innovations'", NBER Working Paper #16443.

Akcigit, U., Baslandze, S., and Stantcheva, S (2015) "Taxation and the International Mobility of Inventors," NBER Working Paper #21024, forthcoming in the *American Economic Review.*

Alvaredo, F., Atkinson, A., Piketty, T. and Saez, E (2014), *The World Top Incomes Database*, http://topincomes.g-mond.parisschoolofeconomics.eu/.

Bakija, J., Cole, A. and Heim, B (2008), "Jobs and Income Growth of Top Earners and the Causes of Changing Income Inequality: Evidence from U.S. Tax Return Data," *Department of Economics Working Papers 2010-22*, Department of Economics, Williams College, revised Jan 2012.

Balkin, D., Markman, G. and Gomez-Mejia, L (2001), "Is CEO Pay in High-Technology Firms Related to Innovation?," *The Academy of Management Journal*, 43 (6). 1118-1129.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., and Van Reenen, J (2015), "The Lifecycle of Inventors," mimeo Harvard.

Black, S, and Devereux, P (2011), "Recent Developments in Intergenerational Mobility," in the *Handbook of Labor Economics, vol. 4*, O. Ashenfelter and D. Card eds., Elsevier, 1487-1541.

Blundell, R., Griffith, R., and Van Reenen, J (1995), "Dynamic Count Data Models of Technological Innovation," *Economic Journal*, 105, 333-344.

Brown, A (2015), "The Richest person in every state, Forbes, http://www.forbes.com/richest-in-each-state/list/#tab:overall

Caballero, R. and Jaffe, A (1993): "How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth," NBER Chapters,in: *NBER Macroeconomics Annual 1993*, Volume 8, pages 15-86 National Bureau of Economic Research, Inc.

Caselli, F (1999), "Technological Revolutions," *American Economic Review*, 89, 78-102.

Chetty, R., Hendren, N., Kline, P., and Saez, E (2014), "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States", *Quarterly Journal of Economics,* 129, 1553-1623.

Depalo, D., and Di Addario, S (2014), "Shedding Light on Inventors' Returns to Patents", *IRLE Working Paper 115-14*, University of California at Berkeley

Deaton, A (2013), *The Great Escape: Health, Wealth, and the Origins of Inequality,* Princeton University Press

Elsby, M., Hobijn, B., and Sahin, A (2013), "The Decline of the US Labor Share", *Brookings Papers on Economic Activity,* Fall 2013, 1-63.

Forbes, K (2000), "A Reassessment of the Relationship between Inequality and Growth", *American Economic Review*, 90, 869-887.

Frank, M (2009), "Inequality and Growth in the United States: Evidence From A New State-Level Panel of Income Inequality Measures", *Economic Inquiry*, 47, 55-68.

Gabaix, X., and Landier, A (2008), "Why Has CEO Pay Increased So Much", *Quarterly Journal of Economics,* 123, 49-100

Galor, O., and Moav, O (2000), "Ability-Biased Technological Transition, Wage Inequality, and Economic Growth", *Quarterly Journal of Economics*, 115, 469-497.

Goldin, C., and Katz, L (2008), *The Race Between Education and Technology,* Harvard University Press

Guvenen, F., Kaplan, G., and Song, J (2014), "The Distribution of Lifetime Incomes in the United States", Meeting Papers 536, Society for Economic Dynamics.

Hall, B., Jaffe, A., and Trajtenberg, M (2001), "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools," *CEPR Discussion Papers 3094, C.E.P.R. Discussion Papers.*

Hall, B., Jaffe, A., and Trajtenberg, M (2005), "Market Value and Patent Citations," *Rand Journal of Economics,* 36 (1), 16-38.

Hassler, J., and Rodriguez Mora, J (2000), "Intelligence, Social Mobility, and Growth", *American Economic Review*, 90, 888-908.

Hémous, D., and Olsen, M (2014), "The Rise of the Machines: Automation, Horizontal Innovation and Income Inequality" *CEPR Discussion Paper DP 10244.*

Jones, C., and Kim, J (2014), "A Schumpeterian Model of Top Income Inequality", mimeo Stanford.

Karabarbounis, L., and Neiman, B (2014), "The Global Decline of the Labor Share", *Quarterly Journal of Economics,* 129, 61-103.

Katz, L., and Murphy, K (1992), "Change in Relative Wages: Supply and Demand Factors", *Quarterly Journal of Economics,* 107, 35-78

Kneer, C (2013), "The Absorption of Talent into Finance: Evidence from U.S. Banking Deregulation," DNB Working Papers 391, Netherlands Central Bank, Research Department.

Kortum, S., and Lerner, J. (1999), "What Is Behind the Recent Surge in Patenting?", *Research Policy*, 28, 1-22.

Krusell, P., Ohanian, L., Rios Rull, V., and Violante, G (2000), "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis", *Econometrica*, 68, 1029-1053.

Lai, R., D'Amour, A., Yu, A., Sun, Y., and Fleming, L (2013), "Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975 - 2010)"

Lerner, J. (1994), "The Importance of Patent Scope: An Empirical Analysis,â€Ý *RAND Journal of Economics*, 1994, 25 (2), 319â€“333.

Moretti, E., and Wilson, D (2014), "State Incentives for Innovation, Star Scientists and Jobs: Evidence from Biotech," *Journal of Urban Economics,* 79, 20-38.

Philippon, T., and Reshef, A (2012), " Wages and Human Capital in the U.S. Finance Industry: 1909-2006," *The Quarterly Journal of Economics,* 127, 1551-1609.

Piketty, T (2014), *Capital in the 21st Century*, Harvard University Press.

Piketty, T., and Saez, E (2003), "Income Inequality in the United States: 1913-1998", *Quarterly Journal of Economics*, 1, 1-39

Piketty, T., Saez, E. and Stantcheva, S (2014), "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities",*American Economic Journal: Economic Policy*, 6, 230-271.

Rajan, R., and Zingales, L (1998), "Financial Dependence and Growth," *American Economic Review*, American Economic Association, 88, 559-86.

Rebelo, S (1991): "Long-Run Policy Analysis and Long-Run Growth," *Journal of Political Economy*, 99, 500-521.

Romer, P (1990), "Endogenous Technical Change", *Journal of Political Economy*, 98, 71-102

Rosen, S (1981), "The Economics of Superstars", *American Economic Review,* 71, 845-858.

Solon, G (1999), "Intergenerational Mobility in the United States", in the *Handbook of Labor Economics, vol.* 3, O. Ashenfelter and D. Card eds., Elsevier, 1761-1800.

Song, J., Price, D., Guvenen, F., and Bloom, N (2015), "Firming Up Inequality", mimeo Stanford University

Toivanen, O., and Vaananen, L (2012), " Returns to Inventors", *Review of Economics and Statistics,* 94, 1173-1190.

Toole, A (2007), "Does Public Scientific Research Complement Private Investment in Research and Development in the Pharmaceutical Industry?", *Journal of Law and Economics*, 50, 81-104.

Wolfers, J (2006), "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results," *American Economic Review*, 96, 1802-20.

Tables used in the main text

Table 1: Descriptive statistics by state in two distinctive years

State	1980		2005		State	1980		2005	
	Innovation	**Top 1%**	**Innovation**	**Top 1%**		**Innovation**	**Top 1%**	**Innovation**	**Top 1%**
AK	66	5.33	134	12.47	**MT**	79	8.02	685	16.35
AL	90	10.01	346	18.48	**NC**	150	9.03	1 878	17.04
AR	61	10.05	231	16.68	**ND**	89	9.62	926	13.23
AZ	381	8.56	2 209	22.66	**NE**	75	9.33	639	15.24
CA	586	9.91	7 082	24.20	**NH**	443	8.48	3 941	18.01
CO	437	9.31	3 165	19.56	**NJ**	1 051	9.83	3 006	20.77
CT	861	12.24	3 395	31.02	**NM**	142	8.90	1 168	15.63
DC	199	14.48	819	23.94	**NV**	310	11.09	2 394	33.30
DE	1 271	10.19	2 311	21.38	**NY**	481	12.08	2 464	30.25
FL	234	12.23	942	31.78	**OH**	449	8.98	1 742	15.86
GA	130	8.95	1 336	19.11	**OK**	529	11.44	878	17.74
HI	58	7.52	551	16.47	**OR**	215	8.25	4 830	16.91
IA	202	8.24	990	12.92	**PA**	492	9.37	1 524	18.71
ID	156	7.68	7 320	18.08	**RI**	246	10.25	1 582	17.36
IL	521	9.63	1 747	21.67	**SC**	158	8.16	564	17.74
IN	368	8.44	1 458	15.52	**SD**	51	8.58	250	16.94
KS	142	10.17	1 328	16.09	**TN**	170	10.09	800	18.76
KY	174	9.69	487	15.76	**TX**	389	12.18	2 267	21.90
LA	139	11.22	293	17.65	**UT**	301	7.79	2 575	18.49
MA	680	10.03	6 794	23.79	**VA**	245	7.97	1 333	17.12
MD	402	8.13	1 998	17.34	**VT**	385	7.97	6 359	16.31
ME	135	8.55	598	15.66	**WA**	245	8.37	9 994	19.69
MI	562	8.91	2 331	16.12	**WI**	303	8.21	1 754	16.48
MN	545	9.31	4 899	18.24	**WV**	187	9.54	205	14.97
MO	185	9.96	896	17.11	**WY**	82	9.00	620	28.52
MS	40	10.48	182	15.81					

Notes: Number of citations within a five-year window per capita and top 1% income share for all 51 states in 1980 and 2005.

Table 2: Descriptive statistics by measures of innovation and top 1% income share in two distinctive years

1980	Mean	p25	p50	p75	Min	Max
Top 1%	9.45	8.37	9.31	10.09	5.33	14.48
Patents	141	71	113	189	27	492
Cit5	312	139	234	443	40	1271
Claims	1564	733	1192	2321	245	5959
Generality	35	17	25	47	5	138
Top5	8	2	5	11	0	40
Top1	1	0	1	2	0	4
2005	Mean	p25	p50	p75	Min	Max
Top 1%	19.07	16.12	17.65	20.77	12.47	33.30
Patents	292	127	228	407	46	898
Cit5	2122	639	1458	2464	134	9994
Claims	5491	2234	4411	763	784	20117
Generality	120	50	89	184	15	344
Top5	12	3	8	14	0	80
Top1	4	1	3	5	0	18

Notes: Summary statistics includes mean, quartiles' thresholds, minimum and maximum for our six measures of innovation and the top 1% income share (relevant variables are defined in Table 3).

Table 3: Variable description and notation

Variable names	**Description**
	Measures of inequality
Top 1%	Share of income own by the richest 1%.
Top 10%	Share of income own by the richest 10%.
Avgtop	Average income share for the percentiles 10 to 2 in the income distribution.
Gini	Gini index of inequality.
G99	Gini index restricted to the bottom 99% of income distribution.
Theil	Theil index of inequality.
Atkinson	Atkinson index of inequality.
	Measures of innovation
Patent	Number of patents granted by the USPTO per inhabitants.
Cit5	Total number of citation received no longer than 5 years after applications per inhabitant.
Claims	Total number of claims associated with patents per inhabitants.
Generality	Total number of patents weighted by the generality index per inhabitants.
Top5	Number of patents in the top 5% most cited per inhabitants.
Top1	Number of patents in the top 1% most cited per inhabitants.
	Measures of social mobility
AM25	Expected percentile of a child at 30 whose parents belonged to the 25^{th} percentile of income distribution in 2000.
AM50	Expected percentile of a child at 30 whose parents belonged to the 50^{th} percentile of income distribution in 2000.
P5-i	Probability for a child at 30 to belong to the 5^{th} quintile of income distribution if parent belonged to the i^{th} quintile, $i \in \{1, 2\}$.
P5	Probability for a child at 30 to belong to the 5^{th} quintile of income distribution if parent belonged to lower quintiles.
	Control variables
Gdppc	Real GDP per capita in US $ (in log).
Popgrowth	Growth of total population.
Sharefinance	Share of the state GDP accounted for by the financial sector. Between 0 and 1.
Unemployment	Unemployment rate. Between 0 and 1.
Gvtsize	Share of the state GDP accounted for by the government sector. Between 0 and 1.
	Additional control variables at the CZ level
Participation Rate	Labor force participation rate.
College per capita	College graduation rate.
School Expenditure	Average expenditures per student in public schools (in log).
Employment Manuf	Share of employed persons 16 and older working in manufacturing.

Notes: Description of relevant variables used in the next tables regressions. Additional variables may be used in specific analysis, in this case they will be explained in the corresponding table description.

Table 4: Top 1% income share and innovation

Dependent variable	Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.019	0.030***	0.025**	0.031***	0.013***	0.008**
	(1.46)	(3.53)	(2.28)	(2.69)	(3.01)	(2.00)
Gdppc	-0.058	-0.091	-0.090	-0.089	-0.078	-0.077
	(-0.99)	(-1.52)	(-1.47)	(-1.47)	(-1.32)	(-1.30)
Popgrowth	0.148	-0.100	-0.117	-0.083	-0.139	-0.134
	(0.16)	(-0.10)	(-0.12)	(-0.08)	(-0.14)	(-0.13)
Sharefinance	0.274	0.433**	0.386**	0.397**	0.391**	0.321*
	(1.41)	(2.24)	(2.01)	(2.08)	(2.04)	(1.72)
Gvtsize	0.119	0.184	0.086	0.101	0.193	0.115
	(0.37)	(0.55)	(0.25)	(0.30)	(0.56)	(0.34)
Unemployment	-0.422	-0.645	-0.595	-0.588	-0.624	-0.563
	(-0.97)	(-1.48)	(-1.36)	(-1.36)	(-1.42)	(-1.31)
R^2	0.908	0.915	0.914	0.915	0.915	0.914
Observations	1785	1632	1632	1632	1632	1632

Notes: The table presents estimates of different measures of innovation on the top 1% income share of state income. We consider different measures of innovation which are all lagged by 2 years and standardized by state population: column (1) uses the number of patents, column (2) uses the number of citations received within a five-year window, column (3) uses the number of claims, column (4) uses the number of patents weighted by their generality index, column (5) uses the number of patents belonging to the top 5% most cited in the year and column (6) uses the number of patents belonging to the top 1% most cited in the year. All these measures as well as the dependent variable are taken in log. Time span: 1977-2011 for column (1) and 1977-2008 for columns (2) to (6). Variable description is given in Table 3.
Panel data OLS regressions with state and year fixed effects. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 5: Innovation and various measures of inequality

Dependent variable	Top 1%	Avgtop	Top 10 %	Overall Gini	G99	Atkinson
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Innovation	0.030***	0.011***	0.019***	-0.001	-0.007*	0.015***
	(3.53)	(2.78)	(4.29)	(-0.17)	(-1.65)	(3.58)
Gdppc	-0.091	-0.054**	-0.053**	0.022	0.029	0.061**
	(-1.52)	(-2.19)	(-2.02)	(1.09)	(1.18)	(2.39)
Popgrowth	-0.100	0.378	0.415	-0.319*	-0.403*	0.374
	(-0.10)	(0.96)	(0.99)	(-1.79)	(-1.78)	(1.32)
Sharefinance	0.433**	0.250***	0.499***	0.100	-0.078	0.304**
	(2.24)	(3.28)	(4.22)	(1.45)	(-1.05)	(2.42)
Gvtsize	0.184	-0.145	-0.692***	0.333***	0.691***	-0.478***
	(0.55)	(-0.86)	(-3.60)	(2.66)	(4.27)	(-3.14)
Unemployment	-0.645	0.017	-0.222	-0.023	0.164	-0.107
	(-1.48)	(0.11)	(-1.24)	(-0.20)	(1.08)	(-0.70)
R^2	0.915	0.430	0.821	0.870	0.750	0.933
Observations	1632	1632	1632	1632	1632	1632

Notes: The table presents estimates of one measure of innovation (citations received within a five-year window per inhabitants) on various measures of inequality: column (1) uses the top 1% income share, column (2) uses the average size of percentiles 2 to 10 in the income distribution, column (3) uses the 10% income share, column (4) uses the gini coefficient, column (5) uses the gini coefficient excluding the first percentile of the income distribution and column (6) uses the Atkinson index with a coefficient of 0.5. Innovation measures have been lagged by 2 years and are taken in log. The dependent variable is also in log in all columns. Time span: 1976-2008. Variable description is given in Table 3.

Panel data OLS regressions with state and year fixed effects. t/z statistics in parentheses, computed with auto-correlation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 6: Top 1% income share and innovation at different lags

Dependent variable	Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Lag of innovation	1 year	2 years	3 years	4 years	5 years	6 years
Innovation	0.038***	0.029***	0.024**	0.016	0.007	-0.007
	(3.59)	(2.87)	(2.29)	(1.59)	(0.68)	(-0.67)
Gdppc	-0.151**	-0.146**	-0.144**	-0.141**	-0.136*	-0.127*
	(-2.15)	(-2.07)	(-2.05)	(-1.97)	(-1.88)	(-1.79)
Popgrowth	0.680	0.701	0.766	0.749	0.740	0.687
	(0.57)	(0.59)	(0.65)	(0.63)	(0.63)	(0.58)
Sharefinance	0.634***	0.582***	0.550***	0.512***	0.480**	0.446**
	(3.19)	(2.92)	(2.78)	(2.63)	(2.49)	(2.33)
Gvtsize	0.201	0.177	0.172	0.137	0.097	0.030
	(0.53)	(0.46)	(0.43)	(0.34)	(0.24)	(0.07)
Unemployment	-0.572	-0.488	-0.432	-0.390	-0.334	-0.279
	(-1.14)	(-0.97)	(-0.87)	(-0.78)	(-0.67)	(-0.57)
R^2	0.878	0.877	0.877	0.877	0.876	0.876
Observations	1428	1428	1428	1428	1428	1428

Notes: The table presents estimates of one measure of innovation (citations received within a five-year window per inhabitants) on the top 1% income share at different lags column (1) uses a one-year lag between the measure of innovation and the dependent variable, column (2) uses two-year lags etc. Both our measure of innovation and the dependent variable are taken in log in all columns. Time span: 1981-2008. Variable description is given in Table 3.

Panel data OLS regressions with state and year fixed effects. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 7: Top 1% income share, innovation and the role of lobbying intensity

Dependent variable	Top 1% Income Share				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Measure of innovation	Cit5	Claims	Generality	Top5	Top1
Innovation	0.142***	0.078*	0.139***	0.082***	0.044***
	(4.13)	(1.81)	(2.98)	(5.23)	(2.75)
Innovation * Lobbying	-0.217***	-0.220***	-0.250***	-0.149***	-0.096***
	(-5.50)	(-4.84)	(-5.38)	(-5.60)	(-3.24)
Lobbying	1.430***	1.822***	1.065***	0.224	0.098
	(4.89)	(4.61)	(4.38)	(1.63)	(0.63)
Gdppc	0.020	0.019	0.010	-0.010	0.006
	(0.26)	(0.23)	(0.13)	(-0.12)	(0.08)
Popgrowth	3.477	4.055*	3.874*	3.732	4.220*
	(1.56)	(1.71)	(1.64)	(1.63)	(1.73)
Sharefinance	0.562	0.506	0.488	0.486	0.556
	(1.40)	(1.27)	(1.22)	(1.20)	(1.46)
Gvtsize	2.829***	2.684***	2.803***	2.635***	2.406***
	(3.07)	(2.77)	(3.01)	(2.81)	(2.59)
Unemployment	-1.197	-0.857	-0.819	-1.290	-1.258
	(-1.33)	(-0.92)	(-0.88)	(-1.44)	(-1.30)
R^2	0.753	0.743	0.745	0.752	0.735
Observations	561	561	561	561	561

Notes: The table presents estimates of different measures of innovation on the top 1% income share of state income. We consider different measures of innovation which are all lagged by 2 years and standardized by state population: column (1) uses the number of citations received within a five-year window, column (2) uses the number of claims, column (3) uses the number of patents weighted by their generality index, column (4) uses the number of patents belonging to the top 5% most cited in the year and column (5) uses the number of patents belonging to the top 1% most cited in the year. All these measures as well as the dependent variable are taken in log. *Lobbying* is measured as the total amount of lobbying in the state divided by population as explained in section 4.4. In each case, we add an interaction terms between our measure of innovation and our measure of lobbying intensity. This dummy is not included in the regression as is would be captured by state fixed effect. Time span: 1998-2008 for all columns. Variable description is given in Table 3.
Panel data OLS regressions with state and year fixed effects. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 8: Top 1% income share and innovation by entrants and incumbents

Dependent variable	Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents			Cit5		
Innovation by Entrants	0.031**		0.031**	0.014**		0.013**
	(2.30)		(2.26)	(2.28)		(2.06)
Innovation by Incumbents		0.011	0.002		0.011	0.008
		(0.87)	(0.16)		(1.37)	(1.05)
Gdppc	-0.107	-0.130*	-0.108	-0.189***	-0.190***	-0.195***
	(-1.53)	(-1.87)	(-1.51)	(-2.74)	(-2.68)	(-2.84)
Popgrowth	0.890	0.718	0.891	-0.022	-0.188	-0.058
	(0.80)	(0.64)	(0.80)	(-0.02)	(-0.13)	(-0.04)
Sharefinance	0.429**	0.488**	0.434**	0.585**	0.721***	0.634**
	(2.24)	(2.54)	(2.19)	(2.42)	(3.01)	(2.53)
Unemployment	-0.303	-0.446	-0.311	-0.291	-0.422	-0.347
	(-0.66)	(-0.93)	(-0.66)	(-0.54)	(-0.75)	(-0.65)
Gvtsize	0.162	0.130	0.162	-0.530	-0.496	-0.486
	(0.43)	(0.34)	(0.44)	(-1.21)	(-1.09)	(-1.10)
R^2	0.880	0.878	0.881	0.828	0.827	0.829
Observations	1479	1479	1479	1224	1224	1224

Notes: The table presents estimates of two different measures of innovation lagged by two years (number of patents and number of citations within a five-year window, denoted Cit3, per inhabitants) on the top 1% income share of state income. We consider two different types of innovation, innovation by entrants as defined by assignees that first patented less than three years ago and innovation by incumbents as defined by assignees that first patented more than three years ago: columns (1) to (3) use the number of patents, columns (4) to (6) use the number of citations within a five-year window. All these measures as well as the dependent variable are taken in log. Time span: 1981-2008 for columns (1) to (3) and 1981-2004 for columns (4) to (6) due to availability of the disambiguation database on assignees and inventors from Lai *et al.* (2013). Variable description is given in Table 3.

Panel data OLS regressions with state and year fixed effects. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 9: Descriptive statistics on the Senate appropriation committee composition.

	Number of years with			Number of years with	
State	1 Senator	2 Senators	**State**	1 Senator	2 Senators
AK	28	0	**MT**	22	0
AL	14	0	**NC**	2	0
AR	28	0	**ND**	25	10
AZ	19	0	**NE**	16	0
CA	14	0	**NH**	32	0
CO	17	0	**NJ**	26	0
CT	12	0	**NM**	36	0
DC	0	0	**NV**	31	1
DE	3	0	**NY**	13	0
FL	21	0	**OH**	6	0
GA	10	0	**OK**	15	0
HI	32	6	**OR**	24	0
IA	19	4	**PA**	35	0
ID	24	0	**RI**	11	0
IL	12	0	**SC**	33	0
IN	9	0	**SD**	17	0
KS	7	0	**TN**	19	0
KY	26	0	**TX**	19	0
LA	32	0	**UT**	27	0
MA	8	0	**VA**	0	0
MD	28	1	**VT**	29	2
ME	3	0	**WA**	21	9
MI	1	0	**WI**	30	8
MN	0	0	**WV**	38	0
MO	29	0	**WY**	7	0
MS	30	8			

Notes: The table gives the number of years between 1970 and 2008 with exactly one (resp. 2) senator seating in the appropriation committee. The exact composition can be found in the appropriation committee official website.

Table 10: Regression of innovation on top 1% income share using instrument based on Appropriation Committee composition in the Senate

Dependent variable	Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.237*	0.172**	0.210**	0.292*	0.129*	0.148*
	(1.85)	(2.03)	(1.97)	(1.81)	(1.93)	(1.81)
Gdppc	-0.200*	-0.212**	-0.251**	-0.253**	-0.150*	-0.207*
	(-1.88)	(-2.21)	(-2.22)	(-2.10)	(-1.81)	(-1.95)
Popgrowth	0.377	-0.011	0.073	0.364	-0.031	0.104
	(0.36)	(-0.01)	(0.07)	(0.32)	(-0.03)	(0.08)
Sharefinance	0.726**	0.936**	0.894**	1.082**	0.998**	0.719*
	(1.98)	(2.28)	(2.21)	(2.09)	(2.12)	(1.82)
Gvtsize	-0.166	0.114	-0.218	-0.071	0.873	0.771
	(-0.44)	(0.24)	(-0.53)	(-0.14)	(1.11)	(0.94)
Unemployment	-1.131	-1.229*	-1.224*	-1.227*	-1.505*	-1.323
	(-1.57)	(-1.79)	(-1.79)	(-1.71)	(-1.79)	(-1.63)
Highways	0.027**	0.030**	0.029**	0.030**	0.031**	0.031**
	(2.18)	(2.35)	(2.24)	(2.07)	(2.28)	(2.08)
Military	0.008	0.011*	0.009*	0.011*	0.008	0.008
	(1.59)	(1.85)	(1.72)	(1.85)	(1.54)	(1.27)
R^2	0.887	0.897	0.897	0.874	0.874	0.834
1^{st} stage F-stat	13.84	14.64	14.43	7.91	10.33	7.28
Observations	1750	1600	1600	1600	1600	1600

Notes: The table presents estimates of different measures of innovation lagged by two years on the top 1% income share of state income: column (1) uses the number of patents, column (2) uses the number of citations received within a five-year window, column (3) uses the number of claims, column (4) uses the number of patents weighted by their generality index, column (5) uses the number of patents belonging to the top 5% most cited in the year and column (6) uses the number of patents belonging to the top 1% most cited in the year. All these measures as well as the dependent variable are taken in log. DC is removed from the sample because it has no senator. Time span: 1977-2011 for column (1) and 1977-2008 for columns (2) to (6). Variable description is given in Table 3.

Panel data IV 2SLS regressions with state and year fixed effects. Innovation is instrumented by the number of senators that seat on the appropriation committee. The lag between the instrument and the endogeneous variable is set to 3 years. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 11: Robustness 1: regression of innovation on top 1% income share using two instruments

Dependent variable	Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.219***	0.132***	0.160***	0.174***	0.108***	0.180***
	(3.47)	(3.65)	(3.51)	(3.49)	(3.52)	(3.12)
lgdppc	-0.208**	-0.165*	-0.179*	-0.150	-0.104	-0.194*
	(-2.12)	(-1.74)	(-1.82)	(-1.56)	(-1.12)	(-1.68)
Highways	0.028***	0.030***	0.028***	0.027**	0.033***	0.041***
	(2.69)	(2.89)	(2.60)	(2.47)	(3.08)	(2.99)
Military	0.007	0.009	0.007	0.006	0.006	0.015*
	(1.25)	(1.50)	(1.25)	(1.08)	(1.05)	(1.80)
Popgrowth	1.796	1.453	1.484	1.447	1.426	1.676
	(1.48)	(1.23)	(1.31)	(1.24)	(1.13)	(1.18)
Sharefinance	0.971***	1.246***	1.157***	1.187***	1.450***	1.593***
	(3.75)	(4.85)	(4.60)	(4.71)	(4.63)	(3.83)
Gvtsize	0.040	0.134	-0.018	0.025	0.892	1.714*
	(0.08)	(0.26)	(-0.04)	(0.05)	(1.36)	(1.80)
Unemployment	-1.153*	-0.992*	-0.931	-0.900	-1.323*	-1.615*
	(-1.82)	(-1.67)	(-1.62)	(-1.57)	(-1.92)	(-1.84)
Spatial Corr	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.797	0.834	0.834	0.832	0.804	0.638
1^{st} stage F-stat	39.56	53.98	51.49	46.94	35.46	11.93
Sargan-Hansen J-stat (p-value)	0.130	0.193	0.0844	0.086	0.418	0.818
Observations	1450	1300	1300	1300	1300	1300

Notes: The table presents estimates of different measures of innovation lagged by two years on the top 1% income share of state income: column (1) uses the number of patents, column (2) uses the number of citations received within a five-year window, column (3) uses the number of claims, column (4) uses the number of patents weighted by their generality index, column (5) uses the number of patents belonging to the top 5% most cited in the year and column (6) uses the number of patents belonging to the top 1% most cited in the year. All these measures as well as the dependent variable are taken in log. DC is removed from the sample because it has no senator. Time span: 1983-2011 for columns (1) 1983-2008 for columns (2) to (6). Variable description is given in Table 3.
Panel data IV 2SLS regressions with state and year fixed effects. Innovation is instrumented by the number of senators that seat on the appropriation committee and by a measure of spillover as described in section6.1. The lag between the first instrument and the endogeneous variable is set to 3 years while the lag between the second instrument and the endogeneous variable is 1 year. Control for spatial correlation involves adding two additional controls for demand shocks as explained in subsection 6.1. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 12: Robustness 2: financial sector, natural resources and taxation

Dependent variable	Top 1% Income Share						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Innovation	0.030***	0.032***	0.030***	0.020**	0.028***	0.030***	0.032***
	(3.42)	(3.64)	(3.56)	(2.14)	(3.30)	(3.49)	(3.64)
Gdppc	-0.085	-0.078	-0.091	-0.092	-0.061	-0.091	-0.097
	(-1.39)	(-1.24)	(-1.52)	(-1.50)	(-1.07)	(-1.52)	(-1.60)
Popgrowth	-0.061	-0.148	-0.101	-0.122	-0.152	-0.104	-0.114
	(-0.06)	(-0.14)	(-0.10)	(-0.12)	(-0.15)	(-0.10)	(-0.11)
Sharefinance	0.485**	0.836***	0.430**	0.415**	0.364*	0.432**	0.432**
	(2.49)	(3.76)	(2.24)	(2.19)	(1.93)	(2.24)	(2.25)
Gvtsize	0.137	0.291	0.165	0.228	0.249	0.175	0.126
	(0.42)	(0.85)	(0.49)	(0.68)	(0.71)	(0.53)	(0.38)
Unemployment	-0.582	-1.111**	-0.639	-0.720*	-0.706	-0.644	-0.710
	(-1.30)	(-2.51)	(-1.47)	(-1.65)	(-1.64)	(-1.48)	(-1.62)
RemunFinance	-0.001						
	(-1.08)						
EFD				0.385***			
				(2.81)			
Oil					-0.003		
					(-0.63)		
Mining					1.208*		
					(1.78)		
Margtax							0.004
							(1.12)
R^2	0.915	0.919	0.915	0.916	0.916	0.915	0.915
Observations	1632	1504	1632	1632	1632	1632	1632

Notes: The table presents estimates of the number of citations received with a five-year window per inhabitants lagged by two years on the top 1% income share of state income: in column (1) we control for average compensation in the financial sector, in column (2), NY, CT, DE and MA (the state with the largest financial sectors) are dropped from the dataset, in column (3), finance-related patents have been removed, in column (4) we control for financial dependence in the state as explained in section 6.2, in column (5) we control for the size of oil and mining sectors, in column (6) oil-related patents have been removed in the count of citations and in column (7) we control for the maximum marginal tax rate. Time Span: 1977-2008. Variables *Oil* and *NaturalRessource* measure the share of oil related and natural resources extraction activities in GDP, variable *RemunFinance* measures the compensation per employee in the financial sector, variable *EFD* measures the financial dependence of innovation and variable *MarginalTax* measures the highest marginal tax rate of labor. Variable description is given in Table 3.
Panel data OLS regressions with state and year fixed effects. Innovation as well as the top 1% income share are taken in log. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 13: Innovation and inequality at the Commuting Zone level

Dependent variable	Top 1%	Top 1%	Gini	Gini	G99	G99
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Innovation	0.047**	0.053**	-0.002	0.022*	-0.018	0.009
	(2.13)	(2.46)	(-0.17)	(1.69)	(-1.22)	(0.68)
Gdppc	0.475**	0.716***	-0.041	0.280***	-0.279**	0.115
	(2.68)	(4.11)	(-0.35)	(3.16)	(-2.25)	(1.40)
Popgrowth	-1.139*	-0.490	-0.648**	-0.221	0.107	-0.096
	(-1.99)	(-1.22)	(-2.01)	(-0.60)	(0.21)	(-0.25)
Gvtsize	-0.002**	-0.001	-0.001*	-0.000	-0.001	0.000
	(-2.13)	(-0.63)	(-1.87)	(-0.11)	(-1.44)	(0.09)
Participation Rate		-0.912***		-1.508***		-1.735***
		(-2.79)		(-6.82)		(-7.16)
School Expenditure		-0.239*		-0.232**		-0.247***
		(-1.92)		(-2.57)		(-2.77)
College per capita		-0.187*		-0.108*		-0.055
		(-1.69)		(-1.82)		(-1.05)
Employment Manuf		-0.262		-0.350**		-0.365**
		(-1.07)		(-2.03)		(-2.10)
R^2	0.173	0.189	0.034	0.228	0.101	0.335
Observations	660	560	670	560	660	560

Notes: The table presents estimates of the number of patents per inhabitants on various measures of inequality at the Commuting Zone (CZ) level: columns (1) and (2) use the top 1% income share of CZ income, columns (3) and (4) use the Gini index and columns (5) and (6) use the Gini index for the bottom 99% of the income distribution. Columns (2), (4) and (6) add additional controls. All innovation and inequalities measures are averaged over the period 1996-2000 and taken in logs. A dummy equal to one if the CZ belongs to an urban area is included but not reported. Variable description is given in Table 3.
Cross section OLS regressions. t/z statistics in parentheses, computed with heteroskedasticity robust standard errors clustered at the state level. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 14: Innovation and social mobility at the Commuting Zone level

Dependent variable	AM25	P1-5	P2-5	AM25	P1-5	P2-5	P5
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Innovation	0.024***	0.108***	0.063***	0.019**	0.073**	0.046*	0.022
	(3.07)	(3.13)	(2.70)	(2.40)	(2.10)	(1.76)	(1.17)
Gdppc	-0.094*	-0.225	-0.204	-0.139***	-0.384*	-0.356**	-0.271**
	(-1.81)	(-1.09)	(-1.48)	(-3.33)	(-1.84)	(-2.39)	(-2.31)
Popgrowth	0.177	0.603	0.711	0.236	0.588	0.731	0.611
	(0.61)	(0.55)	(0.87)	(0.76)	(0.48)	(0.84)	(0.89)
Gvtsize	0.000	0.002	0.001	0.000	-0.000	-0.001	-0.000
	(1.43)	(1.30)	(0.84)	(0.06)	(-0.19)	(-0.77)	(-0.37)
Participation Rate	0.600***	1.356**	1.274**	0.726***	2.067***	1.692***	1.087**
	(3.76)	(2.19)	(2.45)	(4.50)	(3.22)	(3.14)	(2.55)
School Expenditure	0.116**	0.550**	0.349**	0.096*	0.417**	0.298*	0.153
	(2.07)	(2.65)	(2.20)	(1.81)	(2.05)	(1.91)	(1.36)
College per capita				0.081	0.075	0.081	0.119
				(1.52)	(0.35)	(0.49)	(0.98)
Employment Manuf				-0.333***	-1.566***	-1.273***	-0.677***
				(-3.43)	(-4.27)	(-4.18)	(-2.86)
R^2	0.201	0.182	0.163	0.243	0.215	0.211	0.160
Observations	637	645	645	546	546	546	546

Notes: The table presents estimates of the number of patents per inhabitants on various measures of social mobility at the Commuting Zone (CZ) level: columns (1) and (4) use the expected percentile of the income distribution when your parent belongs to the 25^{th} percentile, columns (2) and (5) use the transition probability between the first and the fifth quantiles in the income distribution, columns (3) and (6) use the transition probability between the second and the fifth quantiles in the income distribution, and column (7) uses the probability of reaching the highest quantile in the income distribution when parents did not belong to this quantile. The number of patents per inhabitants is averaged over the period 2006-2010 and social mobility measures are taken when the child is 30 between 2011 and 2012 compared to his parents during the period 1996-2000, all these measures are taken in logs. A dummy equal to one if the CZ belongs to an urban area is included but not reported. Variable description is given in Table 3.
Cross section OLS regressions. t/z statistics in parentheses, computed with heteroskedasticity robust standard errors clustered at the state level. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 15: Innovation and social mobility at the Commuting Zone level. Entrants and incumbents innovation

Dependent variable	AM25	P1-5	P2-5	AM25	P1-5	P2-5	AM25
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Innovation by Entrants	0.016**	0.058**	0.038**				0.018**
	(2.61)	(2.39)	(2.11)				(2.61)
Innovation by Incumbents				0.007	0.032	0.020	-0.006
				(0.87)	(0.97)	(0.75)	(-0.64)
Gdppc	-0.136***	-0.381*	-0.330**	-0.136***	-0.405*	-0.340**	-0.128***
	(-3.08)	(-1.78)	(-2.11)	(-2.96)	(-1.87)	(-2.14)	(-2.83)
Popgrowth	0.287	0.757	0.827	0.272	0.708	0.792	0.290
	(1.00)	(0.66)	(0.98)	(0.92)	(0.61)	(0.93)	(1.02)
Gvtsize	0.000	-0.000	-0.001	0.000	-0.000	-0.001	0.000
	(0.04)	(-0.22)	(-0.80)	(0.08)	(-0.21)	(-0.76)	(0.07)
Participation Rate	0.785***	2.291***	1.815***	0.758***	2.180***	1.743***	0.799***
	(4.61)	(3.44)	(3.25)	(4.48)	(3.30)	(3.14)	(4.71)
School Expenditure	0.109**	0.467**	0.322**	0.102*	0.442**	0.306*	0.111**
	(2.09)	(2.38)	(2.04)	(1.95)	(2.24)	(1.95)	(2.10)
College per capita	0.081*	0.068	0.090	0.075	0.036	0.071	0.084*
	(1.70)	(0.36)	(0.57)	(1.57)	(0.19)	(0.44)	(1.81)
Employment Manuf	-0.312***	-1.508***	-1.212***	-0.366***	-1.705***	-1.341***	-0.307***
	(-3.16)	(-4.12)	(-3.95)	(-3.70)	(-4.54)	(-4.34)	(-3.04)
R^2	0.260	0.233	0.221	0.243	0.217	0.209	0.261
Observations	541	541	541	541	541	541	541

Notes: The table presents estimates of the number of patents per inhabitants on various measures of social mobility at the Commuting Zone (CZ) level. We consider different measures of social mobility: columns (1), (4) and (7) use the expected percentile of the income distribution when your parent belongs to the 25^{th} percentile, columns (2) and (5) use the transition probability between the first and the fifth quantiles in the income distribution, columns (3) and (6) use the transition probability between the second and the fifth quantiles in the income distribution, Innovation has been considered separately whether patents come from entrants (firms that first patented during the period 2006-2010 and incumbents (firms that first patented before 2006). The number of patents per inhabitants is averaged over the period 2006-2010 and social mobility measures are taken when the child is 30 between 2011 and 2012 compared to his parents during the period 1996-2000, all these measures are taken in logs. A dummy equal to one if the CZ belongs to an urban area is included but not reported. Variable description is given in Table 3.
Cross section OLS regressions. t/z statistics in parentheses, computed with heteroskedasticity robust standard errors clustered at the state level. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 16: Innovation and social mobility at the MSA level. Entrants and incumbents innovation and the role of lobbying

Dependent variable	AM25			
	(1)	**(2)**	**(3)**	**(4)**
Innovation by Entrants	0.012	0.028***		
	(1.28)	(2.72)		
Innovation by Incumbents			0.005	0.014
			(0.73)	(1.46)
Gdppc	0.044	0.030	0.046	0.028
	(1.66)	(0.94)	(1.68)	(0.81)
Popgrowth	0.002	0.000	0.003	0.000
	(1.47)	(0.16)	(1.64)	(0.16)
Sharefinance	0.000	-0.003***	0.000	-0.003**
	(0.15)	(-2.82)	(0.40)	(-2.19)
Gvtsize	-0.001	0.001	-0.001	0.001
	(-0.41)	(0.78)	(-0.47)	(0.86)
R^2	0.107	0.079	0.100	0.049
Observations	176	176	176	176

Notes: The table presents estimates of the number of patents per inhabitants on various measures of social mobility at the Metropolitan Statistical Areas (MSA) level: columns (1) and (3) focus on MSA that are below the median in terms of lobbying intensity during the period 2006-2010 and columns (2) and (4) focus on other MSA. The measure of mobility is the expected percentile of the income distribution when your parent belongs to the 25^{th} percentile. Innovation has been considered separately whether patents come from entrants (firms that first patented during the period 2006-2010 and incumbents (firms that first patented before 2006). The number of patents per inhabitants is averaged over the period 2006-2010 and social mobility measures are taken when the child is 30 between 2011 and 2012 compared to his parents during the period 1996-2000, all these measures are taken in logs. Variable description is given in Table 3.
Cross section OLS regressions. t/z statistics in parentheses, computed with heteroskedasticity robust standard errors clustered at the state level. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

A Online Appendix A

A.1 Proofs for subsections 2.2.2 and 2.2.3

Proof of Proposition 2

The only claim we have not formally proved in the text is that $\frac{\partial^2}{\partial\theta_K\partial z}(1-z)\,x_E^* > 0$ (which immediately implies that the positive impact of an increase in R&D productivity on growth, entrepreneurial share and social mobility is attenuated when barriers to entry are high). Differentiating first with respect to θ_E, we get:

$$\frac{\partial\,(1-z)\,x_E^*}{\partial\theta_E} = -\frac{(1-z)\,x_E^*}{\theta_E - \frac{1}{L}(1-z)^2\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)},$$

which is increasing in z since x_E^* and $(1-z)$ both decrease in z and the denominator $\theta_E + \frac{1}{L}(1-z)^2\left[\frac{1}{\eta_H}-\frac{1}{\eta_L}\right]$ increases in z (recall that $\frac{1}{\eta_L}-\frac{1}{\eta_H}>0$). Similarly, differentiating with respect to θ_I gives:

$$\frac{\partial\,(1-z)\,x_E^*}{\partial\theta_I} = \frac{\frac{1}{L}\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)(1-z)^2}{\theta_E - \frac{1}{L}(1-z)^2\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)}\frac{\partial x_I^*}{\partial\theta_I},$$

which is increasing in z since $\frac{\partial x_I^*}{\partial\theta_I}<0$, and $1-z$ and the denominator both decrease in z. This establishes the proposition.

Proof of Proposition 3

Using (5), we rewrite:

$$\frac{w_t}{Y_t} = \frac{1}{L}\left(1-\pi_L-(\pi_H-\pi_L)\left(x_I^*+(1-z)\,x_E^*\right)\right)$$

We then obtain

$$\frac{\partial\,(w_t/Y_t)}{\partial x_I^*} = -\frac{1}{L}(\pi_H-\pi_L) \text{ and } \frac{\partial\,(w_t/Y_t)}{\partial x_E^*} = -\frac{1-z}{L}(\pi_H-\pi_L).$$

Using (14), we get:

$$\frac{\partial rel\_net\_share}{\partial x_I^*} = \left(\frac{1}{2}(\pi_H-\pi_L)\frac{w_t}{Y_t} + \frac{\pi_H-\pi_L}{L}\left(\begin{array}{c}\pi_L+\frac{1}{2}(\pi_H-\pi_L)\,x_I^*\\ +\left(\frac{1}{2}\pi_H-\pi_L\right)(1-z)\,x_E^*\end{array}\right)\right)\left(\frac{Y_t}{w_t}\right)^2\frac{1}{L_t}$$

$$\frac{\partial rel\_net\_share}{\partial x_E^*} = \left(\begin{array}{c} \left(\frac{1}{2}\pi_H + \frac{1}{2}\frac{w_t}{Y_t} - \pi_L\right)(1-z)\frac{w_t}{Y_t} + \\ \frac{(1-z)(\pi_H - \pi_L)}{L} \left(\begin{array}{c} \pi_L + \frac{1}{2}(\pi_H - \pi_L) x_I^* \\ + \left(\frac{1}{2}\pi_H - \pi_L\right)(1-z) x_E^* \end{array} \right) \end{array} \right) \left(\frac{Y_t}{w_t}\right)^2 \frac{1}{L_t}$$

Note that

$$\begin{aligned} A &= \pi_L + \frac{1}{2}(\pi_H - \pi_L) x_I^* + \left(\frac{1}{2}\pi_H - \pi_L\right)(1-z) x_E^* \\ &= \pi_L \left(1 - \frac{1}{2}(1-z) x_E^*\right) + \frac{1}{2}(\pi_H - \pi_L)\left(x_I^* + (1-z) x_E^*\right) \end{aligned}$$

is positive since $(1-z) x_E^* < 1$. Therefore $\frac{\partial rel\_net\_share}{\partial x_I^*} > 0$ and $\frac{\partial rel\_net\_share}{\partial x_E^*} > 0$ if $\frac{1}{2}\pi_H + \frac{1}{2}\frac{w_t}{Y_t} - \pi_L > 0$.

We know that an increase in θ_E has no impact on x_I^* but decreases x_E^*, therefore we get that it reduces the relative net shares whenever $\frac{1}{2}\pi_H + \frac{1}{2}\frac{w_t}{Y_t} - \pi_L > 0$. An increase in θ_I on the other hand affects both x_I^* but also x_E^*, as we have:

$$\frac{\partial x_E^*}{\partial \theta_I} = \frac{\frac{1}{L}(\pi_H - \pi_L)}{\theta_E - \frac{1}{L}(1-z)^2(\pi_H - \pi_L)} \frac{\partial x_I^*}{\partial \theta_I},$$

We can then write

$$\begin{aligned} & \frac{\partial rel\_net\_share}{\partial \theta_I^*} \\ = & \frac{\partial rel\_net\_share}{\partial x_I^*}\frac{\partial x_I}{\partial \theta_I} + \frac{\partial rel\_net\_share}{\partial x_E^*}\frac{\partial x_E}{\partial \theta_E} \\ = & \left(\begin{array}{c} (\pi_H - \pi_L)\frac{w_t}{Y_t}\frac{1}{2}\frac{\theta_E - \frac{1}{L}(1-z)^2\left(\pi_L - \frac{w_t}{Y_t}\right)}{\theta_E - \frac{1}{L}(1-z)^2(\pi_H - \pi_L)} \\ + \frac{(\pi_H - \pi_L)}{L} A \left(1 + \frac{\frac{1}{L}(1-z)^2(\pi_H - \pi_L)}{\theta_E - \frac{1}{L}(1-z)^2(\pi_H - \pi_L)}\right) \end{array} \right) \left(\frac{Y_t}{w_t}\right)^2 \frac{1}{L_t}\frac{\partial x_I^*}{\partial \theta_I} \end{aligned}$$

Note that $x_E^* < 1$, requires $(\pi_H - w)(1-z) < \theta_E$. Moreover as $L > 1$, we must have

$$\theta_E - \frac{1}{L}(1-z)^2\left(\pi_L - \frac{w_t}{Y_t}\right) > \frac{1}{L}(1-z)^2(\pi_H - \pi_L).$$

Hence the relative net share is always decreasing in θ_I.

Finally consider the case where L is large such that $\frac{w_t}{Y_t}$ is small, then we have

$$\frac{w_t}{Y_t} \approx \frac{1}{L}\left(1-\pi_L-(\pi_H-\pi_L)\left(\frac{\pi_H-\pi_L}{\theta_I}+(1-z)\frac{\pi_H}{\theta_E}\right)\right)$$

therefore

$$\begin{aligned}
&\frac{\partial rel_net_share}{\partial x_E^*} \\
\approx\ & \left(\left(\frac{1}{2}\pi_H-\pi_L\right)\frac{w_t}{Y_t}+\frac{(\pi_H-\pi_L)}{L}\left(\begin{array}{c}\pi_L+\frac{1}{2}(\pi_H-\pi_L)\,x_I^* \\ +\left(\frac{1}{2}\pi_H-\pi_L\right)(1-z)\,x_E^*\end{array}\right)\right)\left(\frac{Y_t}{w_t}\right)^2\frac{1-z}{L_t} \\
\approx\ & \left(\begin{array}{c}\left(\frac{1}{2}\pi_H-\pi_L\right)\left(1-\pi_L-(\pi_H-\pi_L)\left(\frac{\pi_H-\pi_L}{\theta_I}+(1-z)\frac{\pi_H}{\theta_E}\right)\right) \\ +(\pi_H-\pi_L)\left(\pi_L+\frac{1}{2}\frac{(\pi_H-\pi_L)^2}{\theta_I}+\left(\frac{1}{2}\pi_H-\pi_L\right)(1-z)\frac{\pi_H}{\theta_E}\right)\end{array}\right)\left(\frac{Y_t}{w_tL_t}\right)^2(1-z) \\
\approx\ & \left(\left(\frac{1}{2}\pi_H-\pi_L\right)(1-\pi_L)+(\pi_H-\pi_L)\,\pi_L+\frac{1}{2}\frac{\pi_L(\pi_H-\pi_L)^2}{\theta_I}\right)\left(\frac{Y_t}{w_tL_t}\right)^2(1-z)
\end{aligned}$$

Then $\left(\frac{1}{2}\pi_H-\pi_L\right)(1-\pi_L)+(\pi_H-\pi_L)\,\pi_L+\frac{1}{2}\frac{\pi_L(\pi_H-\pi_L)^2}{\theta_I}>0$ is a necessary and sufficient condition when L is arbitrarily large under which a decrease in θ_E increases the relative net share.

A.2 Proofs for subsection 2.2.4

From (11), we have: $\frac{\partial x_I^*}{\partial \eta_L}=-\frac{1}{\eta_L^2}\frac{1}{\theta_I}<0$, whereas:

$$\frac{\partial x_E^*}{\partial \eta_L}=(1-z)\frac{\begin{array}{c}[(1-2x_I^*)\left(\theta_E-(1-z)^2\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)\right) \\ -\left(\pi_H-\frac{1}{\eta_L}(1-x_I^*)-\frac{1}{\eta_H}x_I^*\right)(1-z)^2]\end{array}}{\eta_L^2\left(\theta_E-(1-z)^2\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)\right)^2},$$

the sign of which is ambiguous—intuitively a higher η_L decreases incumbent's rate which increases wages but also has a direct negative impact on wages and higher wages in turn lower entrant innovation.

However, when $\theta_E=\theta_I$, the overall effect of a higher η_L on the aggregate innovation rate is negative; more formally:

$$\begin{aligned}
& \frac{\partial x_I^*}{\partial \eta_L}+\frac{\partial x_E^*}{\partial \eta_L} \\
= \quad & -\frac{1}{\eta_L^2}\frac{1}{\theta}+\frac{(1-z)\left(1-x_I^*\right)}{\eta_L^2\left(\theta-(1-z)^2\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)\right)} \\
& -(1-z)\frac{x_I^*\left(\theta-(1-z)^2\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)\right)+\left(\pi_H-\frac{1}{\eta_L}\left(1-x_I^*\right)-\frac{1}{\eta_H}x_I^*\right)(1-z)^2}{\eta_L^2\left(\theta-(1-z)^2\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)\right)^2} \\
= \quad & -\frac{1}{\eta_L^2\left(\theta-(1-z)^2\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)\right)} \\
& \left(\begin{array}{c}
\frac{z}{\theta}\left(\theta+(1-z)\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)\right) \\
+(1-z)\frac{x_I^*\left(\theta-(1-z)^2\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)\right)+\left(\pi_H-\frac{1}{\eta_L}\left(1-x_I^*\right)-\frac{1}{\eta_H}x_I^*\right)(1-z)^2}{\left(\theta-(1-z)^2\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)\right)}
\end{array}\right) \\
< \quad & 0.
\end{aligned}$$

Overall, we therefore have:

$$\frac{\partial entrepreneur\_share_t}{\partial \eta_L}=\frac{1}{\eta_L^2}\left(1-(1-z)\,x_E^*-x_I^*\right)+\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)\frac{\partial}{\partial \eta_L}\left((1-z)\,x_E^*+x_I^*\right),$$

where the second term is dominated by the first term for θ large enough.

A.3 Extensions

A.3.1 Profit sharing between inventor and developer

Here, we assume that once an innovation has been researched, it still needs to be implemented and that this development phase depends on a CEO's effort. Since we are separating the firm owner from the firm manager, we now consider that a firm's owner does not have the outside option of working as a production worker in case her firm does not produce. The economy is still populated by a mass L of workers and a mass 1 of firm owners (who own both the incumbent firm but also the potential entrant firm). For simplicity, the CEO is assumed to be a worker who gets the opportunity to be CEO for a potential entrant or the incumbent in addition to his work as a production workers.

Hence for the owner of an incumbent firm, expected income (net of research spending and CEO wages) is given by:

$$\begin{aligned}\widetilde{\Pi}^{inc}\left(x_I, e_I, R_{I,H}, R_{I,L}\right) &= e_I x_I \left(\pi_H - R_{I,H}\right) Y_t + \left(1 - e_I x_I - (1-z)\, e_E^* x_E^*\right) \pi_L Y_t \\ &\quad - \left(1 - e_I\right) x_I R_{I,L} Y_t - \theta_I \frac{x_I^2}{2} Y_t,\end{aligned}$$

where e_I denotes the likelihood that the CEO succeeds in ensuring that the company implements the new technology—and similarly e_E^* is the equilibrium likelihood that the CEO of an entrant company manages to set-up a new firm. $R_{I,H}Y_t$ is the income that the CEO obtains in case of a success, and $R_{I,L}Y_t$, his income if he fails.

To obtain a success rate e_I, a CEO has to incur a utility effort cost $\psi \frac{e_I^2}{2} Y_t$. The CEOs outside option is 0 (we assume that he can always reject a negative payment). A CEO of an incumbent firm will then solve the following program:

$$\underset{e_I}{Max} \left\{ e_I R_{I,H} Y_t + \left(1 - e_I\right) R_{I,L} - \psi \frac{e_I^2}{2} Y_t \right\}.$$

We then obtain that the constraint $R_{I,L} \geq 0$ will bind. As a result the CEO will choose a success probability:

$$e_I^* = R_{I,H}^* / \psi.$$

This implies that the firm's owner will decide on a payment

$$R_{I,H}^* = \left(\pi_H - \pi_L\right)/2.$$

Therefore, in case of a success, the CEO obtains half of the gains from innovation.

Similarly for an entrant firm owner, we find that her expected income is given by:

$$\widetilde{\Pi}^{ent}\left(x_E, e_E, R_{E,H}, R_{E.L}\right) = (1-z)\, e_E x_E \left(\pi_H - R_{E,H}\right) Y_t - (1-z)\, x_E \left(1 - e_E\right) R_{E,L} Y_t - \theta_E \frac{x_E^2}{2} Y_t.$$

e_E is now the likelihood that the CEO succeeds in setting up a new firm (here we assumed that the CEO effort is undertaken after the innovation has been potentially blocked, this is without loss of generality). As above the constraint that $R_{E,L} = 0$ binds must be satisfied. We then obtain that $e_E^* = R_{E,H}^* / \psi$ as before, which now leads to

$$R_{E,H}^* = \pi_H / 2.$$

Here as well the CEO gets half of the gains from innovation in case of success.[50]

[50]The gains from an innovation for the owner of an entrant firm is $\pi_H Y_t$, while it was $\pi_H Y_t - w_t$ when she

We obtain that as a share of gross output, CEOs income is given by

$$CEO\_share = x_I^* e_I^* R_{I,H} + (1-z)\, x_E^* e_E^* R_{E.H}^* = \frac{1}{\theta_I}\frac{(\pi_H - \pi_L)^4}{16\psi^2} + \frac{(1-z)^2}{\theta_E}\frac{\pi_H^4}{16\psi^2}.$$

Therefore it decreases with both entrant and incumbent innovation costs. As long as the labor force is large enough, top income earners will be the owners and the CEO. As a share of gross output, their joint income (net of innovation costs) will be given by:

$$Top\_share = \pi_H \mu^* + \pi_L (1-\mu^*) - \frac{\theta_E x_E^2}{2} - \frac{\theta_I x_I^2}{2}, \tag{16}$$

where the share of high-mark up sectors satisfies:

$$\mu^* = x_I^* e_I^* + (1-z)\, x_E^* e_E^*.$$

It is then straightforward to show that this top share decreases with the incumbent innovation costs θ_I, whereas the labor share increases with both entrant and incumbent innovation costs. Furthermore, a decrease in entrant innovation cost θ_E shifts income towards top earners relative to workers (i.e. it increases $Top\_share/wage\_share$) if and only if $3\pi_H - 4\pi_L + \pi_L\pi_H + \pi_L\frac{(\pi_H-\pi_L)^4}{8\theta_I\psi^2} > 0$, which is satisfied if profits of innovative firms are large enough relative to the non-innovative ones. Indeed, entrant innovation can potentially reduce the owner share for the same reasons as above. This establishes:

Proposition 4 *A reduction in incumbents innovation costs favors top income earners. A reduction in entrant's innovation costs favors top income earners if and only if* $3\pi_H - 4\pi_L + \pi_L\pi_H + \pi_L\frac{(\pi_H-\pi_L)^4}{8\theta_I\psi^2} > 0$.

Proof. Solving for the innovation decision we obtain that incumbents invest:

$$x_I^* = \frac{1}{\theta_I}\frac{(\pi_H-\pi_L)^2}{4\psi} = \frac{1}{4\psi\theta_I}\left(\frac{1}{\eta_L} - \frac{1}{\eta_H}\right)^2.$$

Entrants invest

$$x_E^* = \frac{1-z}{4\psi\theta_E}\pi_H^2 = \frac{1-z}{4\psi\theta_E}\left(1 - \frac{1}{\eta_H}\right)^2.$$

We can then express the share of high mark-up sector as:

$$\mu^* = \frac{1}{\theta_I}\frac{(\pi_H-\pi_L)^3}{8\psi^2} + \frac{(1-z)^2}{8\psi^2\theta_E}\pi_H^3.$$

had the outside option of becoming a worker.

Since the wage share is given by

$$\begin{aligned}\frac{w_t L}{Y_t} &= 1-\pi_L-(\pi_H-\pi_L)\,\mu^* \\ &= 1-\pi_L-(\pi_H-\pi_L)\left(\frac{1}{\theta_I}\frac{(\pi_H-\pi_L)^3}{8\psi^2}+\frac{(1-z)^2}{8\psi^2\theta_E}\pi_H^3\right),\end{aligned}$$

both innovation costs increase the labor share of gross output. The top earners share (using (16) and the values for the innovation rates) can then be expressed as:

$$\begin{aligned}Top\_share &= 1-\frac{w_t L}{Y_t}-\left(\frac{(\pi_H-\pi_L)^4}{32\theta_I\psi^2}+\frac{(1-z)^2\,\pi_H^4}{32\theta_E\psi^2}\right), \\ &= \pi_L+\frac{3\,(\pi_H-\pi_L)^4}{32\theta_I\psi^2}+\frac{(1-z)^2\,\pi_H^3\,(3\pi_H-4\pi_L)}{32\psi^2\theta_E}.\end{aligned}$$

Hence we get that Top_share is decreasing in θ_I. Further, we get that

$$\frac{\partial}{\partial\theta_E}\left(\frac{Top\_share}{(w_tL/Y_t)}\right)=-\frac{(1-z)^2\,\pi_H^3}{32\psi^2\theta_E^2}\left(\frac{Y_t}{w_tL}\right)^2\left(3\pi_H-4\pi_L+\pi_L\pi_H+\pi_L\frac{(\pi_H-\pi_L)^4}{8\theta_I\psi^2}\right)$$

Hence an increase in θ_E shifts income towards workers to the detriment of the top earners if $3\pi_H-4\pi_L+\pi_L\pi_H+\pi_L\frac{(\pi_H-\pi_L)^4}{8\theta_I\psi^2}>0$ (which is satisfied if π_H/π_L is large enough). ■

A.3.2 Profit sharing between firm owner and inventor

To distinguish between the firm owner and the innovator we now consider that the set of potential firm owners is given (i.e. there is a mass 1 of capitalists who inherit incumbent firms and can each set up an entrant firm), while innovators are drawn from the population. There is a mass L of potential workers. Workers are identical when in production but differ in the quantity of human capital they can produce in innovation (each worker can produce h units of human capital and h is distributed uniformly over $\left[0,\overline{h}\right]$).

To innovate with probability x an incumbent firm needs to hire $\theta e^2/2$ units of human capital. Similarly an entrant firm needs to hire $\theta e^2/2$ units of human capital.[51] Denoting by v the price of 1 unit of innovative human capital normalized by Y_t, we obtain that there will be a threshold $\widehat{h}$, such that individuals whose h is below $\widehat{h}$ will be production workers and

[51]We assume that the innovation cost is the same for entrants and incumbents. Without this assumption a reduction in entrant's cost could lead to a reduction in overall innovation through its impact on the price of human capital for some extreme parameter assumptions.

those above will be innovators. That threshold obeys

$$\frac{w}{Y} = v\widehat{h}. \tag{17}$$

Solving for the profit maximization problem, we find the optimal innovation rates as:

$$x_I^* = \frac{\pi_H - \pi_L}{\theta v} \text{ and } x_E^* = \pi_H \frac{1-z}{\theta v}, \tag{18}$$

for the incumbent and the entrant respectively. These rates are similar to those in the baseline model, except that they depend on the wage rate v and that the entrant rate does not depend on w (since a firm owner does not have the possibility to become a worker if he fails).

Market clearing for human capital implies:

$$\begin{aligned} \theta \left(\frac{x_I^{*2}}{2} + \frac{x_E^{*2}}{2} \right) &= L \int_{\widehat{h}}^{\overline{h}} h dh \Leftrightarrow \\ (\pi_H - \pi_L)^2 + \pi_H^2 (1-z)^2 &= \theta v^2 L \frac{\overline{h}^2 - \widehat{h}^2}{\overline{h}}. \end{aligned} \tag{19}$$

This equation establishes the demand for innovative human capital as a function of the wage rate and the cost of innovation.

The supply-side equation can be determined by combining (17) with the production labor share equation:

$$\frac{wL\widehat{h}}{Y\overline{h}} = \frac{\mu}{\eta_H} + \frac{1-\mu}{\eta_L},$$

as $L\widehat{h}$ is the labor force in production. We then obtain:

$$vL\frac{\widehat{h}^2}{\overline{h}} = 1 - \pi_L + \frac{\pi_L - \pi_H}{\theta v} \left(\pi_H - \pi_L + \pi_H (1-z)^2 \right). \tag{20}$$

Plugging (20) into (19), we obtain that the wage rate for innovative human capital is uniquely defined by:

$$vL\overline{h} = 1 - \pi_L + \pi_L \pi_H \frac{(1-z)^2}{\theta v}. \tag{21}$$

Hence v is decreasing in θ (i.e. the lower is the cost of innovation, the higher is the level of wage per unit of human capital).

As shown below, a decrease in the innovation cost boosts innovation both by entrants and incumbents. In addition, the threshold $\widehat{h}$ decreases, so that when innovation costs go down, more workers end up working as innovators.

Two measures of inequality can be derived here: the share of income going to the firm owners (here we implicitly assume that firm ownership is concentrated at the top of the income distribution) and a measure of top labor income inequality.

The income share of innovators can be derived as:

$$Innov\_share = \int_{\widehat{h}}^{\overline{h}} vLhdh = vL\left(\overline{h}^2 - \widehat{h}^2\right) / \left(2\overline{h}\right). \tag{22}$$

One can show that this expression is decreasing in θ (hence lower innovation costs increase the share of income going to innovators).

We show below that the owner share of GDP must satisfy:

$$\begin{aligned} Owner\_share &= \pi_L (1-\mu) + \pi_H \mu - Innov\_share \\ &= \pi_L + \frac{1}{2\theta v}\left((\pi_H - \pi_L)^2 + (\pi_H - 2\pi_L)\pi_H (1-z)^2\right). \end{aligned} \tag{23}$$

Hence a reduction in innovation costs will increase the owner share of income as long as $(\pi_H - \pi_L)^2 + (\pi_H - 2\pi_L)\pi_H(1-z)^2 > 0$ (the intuition is still that entrant innovations may decrease overall owner's net share of income by suppressing the rents of an incumbent). If firms' owner are disproportionately concentrated in the top of the income distribution, this predicts that a reduction in innovation will increase top income inequality.

The share of labor income going to the individuals above some ratio $\widetilde{h}/\overline{h}$ can be expressed as

$$\begin{aligned} TopLincome(\widetilde{h}) &= \frac{\int_{\widetilde{h}}^{\overline{h}} vhdh}{\frac{w}{Y}\frac{\widehat{h}}{\overline{h}} + \int_{\widehat{h}}^{\overline{h}} vhdh} = \frac{\overline{h}^2 - \widetilde{h}^2}{\widehat{h}^2 + \overline{h}^2} \text{ if } \widetilde{h} \geq \widehat{h} \\ &= 1 - \frac{\frac{w}{Y}\frac{\widetilde{h}}{\overline{h}}}{\frac{w}{Y}\frac{\widehat{h}}{\overline{h}} + \int_{\widehat{h}}^{\overline{h}} vhdh} = 1 - \frac{2\widehat{h}\widetilde{h}}{\widehat{h}^2 + \overline{h}^2} \text{ if } \widetilde{h} \leq \widehat{h}. \end{aligned}$$

In both cases, $TopLincome$ is decreasing in $\widehat{h}$ and therefore also in innovation costs. One can then prove the following proposition.

Proposition 5 *A reduction in innovation costs leads to an increase in innovation, an increase in top labor income inequality and an increase in the owners' share of income if* $(\pi_H - \pi_L)^2 + (\pi_H - 2\pi_L)\pi_H(1-z)^2 > 0$.

Proof: Using (21) we have:

$$\frac{dv}{d\theta} = \frac{v}{\theta} \frac{-\pi_L \pi_H \frac{(1-z)^2}{\theta v^2}}{L\overline{h} + \pi_L \pi_H \frac{(1-z)^2}{\theta v^2}}.$$

Hence we get:

$$\frac{d(\theta v)}{d\theta} = v \frac{L\overline{h}}{L\overline{h} + \pi_L \pi_H \frac{(1-z)^2}{\theta v^2}} > 0.$$

Using (18) we then obtain that both entrant innovation x^* and incumbent innovation x_I^* decrease with θ. Differentiating (19) we get:

$$\begin{aligned} \frac{d\widehat{h}}{d\theta} &= \frac{\overline{h}^2 - \widehat{h}^2}{2\theta} \left(1 + 2\frac{\theta}{v}\frac{dv}{d\theta}\right) \\ &= \frac{\overline{h}^2 - \widehat{h}^2}{2\theta} \frac{L\overline{h} - \pi_L \pi_H \frac{(1-z)^2}{\theta v^2}}{L\overline{h} + \pi_L \pi_H \frac{(1-z)^2}{\theta v^2}} \\ &= \frac{\overline{h}^2 - \widehat{h}^2}{L\overline{h} + \pi_L \pi_H \frac{(1-z)^2}{\theta v^2}} \frac{1-\pi_L}{2\theta v} > 0, \end{aligned}$$

where we used (21) to obtain the latter equality.

Using (19) in (22), we obtain that the share of income that goes to innovators is given by:

$$Innov\_share = \frac{(\pi_H - \pi_L)^2 + \pi_H^2 (1-z)^2}{2\theta v},$$

which is decreasing in θ since θv is increasing in θ.

To compute the owner share we use the previous equation and (18) in (23) to obtain:

$$\begin{aligned} Owner\_share &= \pi_L + (\pi_H - \pi_L)(x_I^* + (1-z)x_E^*) - Innov\_share \\ &= \pi_L + (\pi_H - \pi_L)\left(\frac{\pi_H - \pi_L}{\theta v} + (1-z)\pi_H \frac{1-z}{\theta v}\right) - \frac{(\pi_H - \pi_L)^2 + \pi_H^2 (1-z)^2}{2\theta v} \\ &= \pi_L + \frac{1}{2\theta v}\left((\pi_H - \pi_L)^2 + (\pi_H - 2\pi_L)\pi_H (1-z)^2\right). \end{aligned}$$

Therefore the owner share is increasing in θ if and only if $(\pi_H - \pi_L)^2 + (\pi_H - 2\pi_L)\pi_H (1-z)^2 > 0$, which establishes the Proposition.

B Online Appendix B: Additional tables

Table B1: Top 0.1% and Top 0.01% income share and innovation from incumbents and entrants

Dependent variable	Top 0.1% Income Share			Top 0.01% Income Share		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Innovation by Entrants	0.017**		0.016**	0.016		0.016
	(2.12)		(1.98)	(1.45)		(1.39)
Innovation by Incumbents		0.035***	0.035***		0.047***	0.048***
		(3.39)	(3.44)		(3.39)	(3.48)
Gdppc	-0.240***	-0.215***	-0.229***	-0.358***	-0.322***	-0.336***
	(-3.31)	(-3.07)	(-3.22)	(-3.43)	(-3.18)	(-3.28)
Popgrowth	4.006***	4.159***	4.058***	5.446***	5.694***	5.554***
	(3.06)	(3.35)	(3.24)	(2.93)	(3.23)	(3.13)
Sharefinance	1.149***	1.340***	1.394***	1.222**	1.491**	1.569***
	(2.81)	(2.99)	(3.41)	(2.12)	(2.34)	(2.72)
Gvtsize	-2.616***	-2.426***	-2.312***	-3.491***	-3.177***	-3.049***
	(-3.93)	(-3.74)	(-3.52)	(-3.75)	(-3.49)	(-3.32)
Unemployment	-0.328	-0.576	-0.607	-0.731	-1.076	-1.125
	(-0.54)	(-0.98)	(-1.03)	(-0.88)	(-1.34)	(-1.40)
R^2	0.781	0.784	0.785	0.724	0.729	0.730
Observations	1224	1224	1224	1224	1224	1224

Notes: The table presents estimates of the number of citations received within a five-year window per inhabitants on the top 0.1% and top 0.01% income share of state income. We consider two different types of innovation, innovation by entrants as defined by assignees that first patented less than 3 years ago and innovation by incumbents as defined by assignees that first patented more than 3 years ago: columns (1) to (3) use the top 0.1% income share and columns (4) to (6) use the top 0.01% income share. Both innovation measures and dependent variables are taken in log. Time span: 1980-2004 due to availability of the disambiguition database on assignees and inventors from Lai *et al.* (2013). Variable description is given in Table 3.
Panel data OLS regressions with state and year fixed effects. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B2: Top 1% income share and innovation with clustered standard errors

Dependent variable	Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.019	0.030***	0.025*	0.031**	0.013**	0.008*
	(0.95)	(2.71)	(1.78)	(2.08)	(2.45)	(1.75)
Gdppc	-0.058	-0.091*	-0.090*	-0.089*	-0.078*	-0.077
	(-1.24)	(-1.89)	(-1.74)	(-1.75)	(-1.65)	(-1.60)
Popgrowth	0.148	-0.100	-0.117	-0.083	-0.139	-0.134
	(0.12)	(-0.07)	(-0.08)	(-0.06)	(-0.09)	(-0.09)
Sharefinance	0.274	0.433	0.386	0.397	0.391	0.321
	(0.73)	(1.26)	(1.11)	(1.15)	(1.14)	(0.96)
Gvtsize	0.119	0.184	0.086	0.101	0.193	0.115
	(0.22)	(0.33)	(0.16)	(0.19)	(0.34)	(0.21)
Unemployment	-0.422	-0.645	-0.595	-0.588	-0.624	-0.563
	(-0.64)	(-0.86)	(-0.79)	(-0.80)	(-0.83)	(-0.76)
R^2	0.908	0.915	0.914	0.915	0.915	0.914
Observations	1785	1632	1632	1632	1632	1632

Notes: The table presents estimates of different measures of innovation on the top 1% income share of state income. We consider different measures of innovation which are all lagged by 2 years and standardized by state population: column (1) uses the number of patents, column (2) uses the number of citations received within a five-year window, column (3) uses the number of claims, column (4) uses the number of patent weighted by their generality index, column (5) uses the number of patents belonging to the top 5% most cited in the year and column (6) uses the number of patents belonging to the top 1% most cited in the year. All these measures as well as the dependent variable are taken in log. Time span: 1976-2011 for column (1) and 1976-2008 for columns (2) to (6). Variable description is given in Table 3.
Panel data OLS regressions with state and year fixed effects. t/z statistics in parentheses, computed with heteroskedasticity robust standard errors clustered at the state level. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B3: INNOVATION AND VARIOUS MEASURES OF INEQUALITY BASED ON DIFFERENT INCOME SHARES

Dependent variable	Top 10%	Top 5%	Top 1%	Top 0.5%	Top 0.1%	Top 0.01%
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Innovation	0.019***	0.021***	0.030***	0.044***	0.052***	0.065***
	(4.29)	(3.97)	(3.53)	(4.61)	(4.12)	(3.56)
Gdppc	-0.053**	-0.038	-0.091	-0.105**	-0.121*	-0.175*
	(-2.02)	(-1.18)	(-1.52)	(-2.06)	(-1.88)	(-1.89)
Popgrowth	0.415	0.407	-0.100	0.979	1.459	2.198
	(0.99)	(0.89)	(-0.10)	(1.19)	(1.32)	(1.41)
Sharefinance	0.499***	0.489***	0.433**	0.823***	0.544	0.171
	(4.22)	(3.07)	(2.24)	(2.74)	(1.22)	(0.27)
Gvtsize	-0.692***	-0.805***	0.184	-1.684***	-1.973***	-2.241***
	(-3.60)	(-3.68)	(0.55)	(-4.67)	(-4.21)	(-3.37)
Unemployment	-0.222	-0.554***	-0.645	-0.948**	-1.135**	-1.645**
	(-1.24)	(-2.64)	(-1.48)	(-2.41)	(-2.13)	(-2.17)
R^2	0.821	0.873	0.915	0.892	0.892	0.865
Observations	1632	1632	1632	1632	1632	1632

Notes: The table presents estimates of one measure of innovation (citations received within a five-year window per inhabitants) on various measures of inequality: column (1) uses the top 10% income share, column (2) uses the top 5%, column (3) uses the top 1% income share, column (4) uses the top 0.5% income share, column (5) uses the top 0.1% income share and column (6) uses the top 0.01% income share. The innovation measure has been lagged by 2 years and is taken in log. The dependent variable is also in log in all columns. Time span: 1976-2008. Variable description is given in Table 3.
Panel data OLS regressions with state and year fixed effects. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B4: Top 1% income share and innovation by entrants and incumbents - alternative definition of entrants

Dependent variable	Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation		Patents			Cit5	
Innovation by Entrants	0.034**		0.035**	0.013*		0.010
	(2.39)		(2.39)	(1.86)		(1.49)
Innovation by Incumbents		0.009	-0.002		0.012	0.008
		(0.74)	(-0.16)		(1.49)	(1.01)
Gdppc	-0.108	-0.129*	-0.107	-0.176***	-0.192***	-0.183***
	(-1.54)	(-1.85)	(-1.50)	(-2.58)	(-2.72)	(-2.70)
Popgrowth	0.898	0.717	0.896	0.075	-0.166	0.059
	(0.81)	(0.64)	(0.81)	(0.06)	(-0.11)	(0.04)
Sharefinance	0.428**	0.485**	0.423**	0.552**	0.721***	0.591**
	(2.25)	(2.52)	(2.14)	(2.29)	(3.01)	(2.36)
Gvtsize	0.170	0.128	0.170	-0.529	-0.505	-0.507
	(0.46)	(0.33)	(0.46)	(-1.21)	(-1.12)	(-1.16)
Unemployment	-0.303	-0.440	-0.295	-0.242	-0.417	-0.280
	(-0.66)	(-0.92)	(-0.63)	(-0.46)	(-0.74)	(-0.54)
R^2	0.881	0.878	0.881	0.829	0.828	0.830
Observations	1479	1479	1479	1224	1224	1224

Notes: The table presents estimates of two different measures of innovation lagged by two years (number of patents and number of citations within a five-year window per inhabitants) on the top 1% income share of state income. We consider two different types of innovation, innovation by entrants as defined by assignees that first patented less than 5 years ago and innovation by incumbents as defined by assignees that first patented more than 5 years ago: columns (1) to (3) use the number of patents, columns (4) to (6) use the number of citations within a five-year window. All these measures as well as the dependent variable are taken in log. Time span: 1981-2008 for columns (1) to (3) and 1981-2004 for columns (4) to (6) due to availability of the disambiguition database on assignees and inventors from Lai *et al.* (2013). Variable description is given in Table 3.
Panel data OLS regressions with state and year fixed effects. t/z statistics in parentheses, computed with auto-correlation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B5: First stage and reduced form regressions

Dependent variable	Cit5	Top 1%	Cit5	Top 1%	Cit5	Top 1%
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Appropriation Committee	0.099***	0.017**			0.041*	0.018**
	(3.88)	(2.26)			(1.67)	(2.02)
Spillover			4.707***	0.573***	4.594***	0.554***
			(10.26)	(3.73)	(9.92)	(3.51)
Gdppc	0.719***	-0.089	0.959***	-0.007	1.050***	-0.019
	(2.77)	(-1.51)	(3.04)	(-0.10)	(3.32)	(-0.28)
Popgrowth	-0.244	-0.053	3.970	1.834*	4.506*	2.123*
	(-0.10)	(-0.05)	(1.61)	(1.76)	(1.79)	(1.95)
Sharefinance	-4.004***	0.248	-2.377***	0.970***	-2.289***	0.926***
	(-5.43)	(1.32)	(-2.89)	(5.00)	(-2.68)	(4.70)
Gvtsize	-3.183**	-0.433	0.690	0.792	2.796	0.454
	(-2.00)	(-1.14)	(0.37)	(1.54)	(1.44)	(0.81)
Unemployment	4.400***	-0.473	7.730***	0.036	7.816***	0.036
	(3.65)	(-1.06)	(6.36)	(0.08)	(6.34)	(0.08)
Highways	-0.107***	0.012				0.021**
	(-3.28)	(1.23)				(2.13)
Military	-0.042***	0.004			-0.037***	0.004
	(-4.43)	(0.92)			(-3.35)	(0.70)
Spatial Corr	No	No	Yes	Yes	Yes	Yes
R^2	0.879	0.918	0.880	0.851	0.883	0.858
Observations	1600	1600	1326	1326	1300	1300

Notes: The table presents the regressions results of our instruments on the innovation variable (measured by the number of citations received within a five-year window) (columns (1), (3) and (5)) and the results of our instruments directly on the dependent variable (the share of income held by the richest 1%) in other columns. Columns (1) and (2) use the state number of senators with a seat on the Senate appropriation committee, columns (3) and (4) use the spillover instruments at three different lags and columns (5) and (6) use all instruments. The lags between the depend variable and the instruments are set to match the corresponding 2 stage regressions: 3 years for column (1), 5 years for column (2), 1 year for columns (3), 3 years for column (4), 3 and 1 years for column (5) and 5 and 3 years for column (6). DC is removed from the sample in columns (1), (2), (5) and (6) because it has no senators. All the dependent variables and the spillover instrument are taken in log. Time Span: 1977-2008 for columns (1) and (2) and 1983-2008 for columns (3) to (6). Variable description is given in Table 3.
Panel data OLS regressions with state and year fixed effects. Innovation as well as the top 1% income share are taken in log. Control for spatial correlation involves adding two additional controls for demand shocks as explained in subsection 6.1. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B6: Innovation and various measures of inequality - IV results

Dependent Variable	Top 1%	Avgtop	Top 10 %	Overall Gini	G99	Atkinson
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Innovation	0.172**	0.034	0.054	-0.010	-0.027	0.045
	(2.03)	(1.22)	(1.32)	(-0.41)	(-0.90)	(1.29)
Gdppc	-0.212**	-0.067**	-0.079**	0.027	0.041	0.037
	(-2.21)	(-2.14)	(-2.06)	(1.04)	(1.25)	(1.01)
Popgrowth	-0.011	0.486	0.425	-0.325*	-0.374	0.361
	(-0.01)	(1.14)	(0.90)	(-1.76)	(-1.61)	(1.19)
Sharefinance	0.936**	0.279**	0.558***	0.058	-0.142	0.429**
	(2.28)	(2.01)	(2.88)	(0.50)	(-1.01)	(2.28)
Gvtsize	0.114	-0.168	-0.804***	0.402**	0.810***	-0.303
	(0.24)	(-0.91)	(-3.75)	(2.30)	(3.72)	(-1.50)
Unemployment	-1.229*	-0.075	-0.342	-0.071	0.122	-0.297
	(-1.79)	(-0.35)	(-1.31)	(-0.43)	(0.56)	(-1.34)
Highways	0.030**	0.003	0.011	0.011***	0.011**	0.011
	(2.35)	(0.71)	(1.58)	(2.99)	(2.43)	(1.53)
Military	0.011*	-0.004**	-0.001	-0.004**	-0.005**	-0.000
	(1.85)	(-2.03)	(-0.53)	(-2.05)	(-2.36)	(-0.22)
R^2	0.897	0.438	0.819	0.873	0.749	0.930
1^{st} stage F-stat	14.64	14.64	14.64	14.64	14.64	14.64
Observations	1600	1600	1600	1600	1600	1600

Notes: The table presents estimates of one measure of innovation (citations received within a five-year window per inhabitants) on various measures of inequality: column (1) uses the top 1% income share, column (2) uses the average size of percentiles 2 to 10 in the income distribution, column (3) uses the 10% income share, column (4) uses the gini coefficient, column (5) uses the gini coefficient excluding the first percentile of the income distribution and column (6) uses the Atkinson index with a coefficient of 0.5. The innovation measure has been lagged by 2 years and is taken in log. The dependent variable is also in log in all columns. Time span: 1977-2008. Variable description is given in Table 3.

Panel data IV 2SLS regressions with state and year fixed effects. Innovation is instrumented by the number of senators that seat on the appropriation committee. DC is removed from the sample in all columns because it has no senators. The lag between the instrument and the endogeneous variable is set to 3 years. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B7: Top 1% income share and innovation at different lags - IV results

Dependent variable	Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Lag of innovation	1 year	2 years	3 years	4 years	5 years	6 years
Innovation	0.200**	0.209**	0.194**	0.164*	0.115	0.082
	(2.11)	(2.20)	(2.05)	(1.91)	(1.45)	(0.94)
Gdppc	-0.281**	-0.284**	-0.271**	-0.255**	-0.246**	-0.183*
	(-2.53)	(-2.52)	(-2.41)	(-2.33)	(-2.08)	(-1.79)
Popgrowth	0.453	0.533	1.127	1.007	1.315	1.355
	(0.34)	(0.40)	(0.83)	(0.77)	(1.00)	(1.08)
sharefinance	1.344***	1.265***	1.129***	0.922***	0.757**	0.628**
	(2.78)	(2.83)	(2.69)	(2.63)	(2.41)	(2.24)
Gvtsize	0.062	0.218	0.513	0.377	0.252	0.260
	(0.13)	(0.43)	(0.89)	(0.69)	(0.49)	(0.46)
Unemployment	-1.649*	-1.620*	-1.387*	-1.226	-0.864	-0.538
	(-1.76)	(-1.81)	(-1.68)	(-1.60)	(-1.28)	(-0.87)
Highways	0.032**	0.033**	0.028**	0.023*	0.013	0.002
	(2.33)	(2.34)	(2.10)	(1.78)	(1.04)	(0.17)
Military	0.015*	0.012*	0.008	0.007	0.003	-0.000
	(1.96)	(1.66)	(1.30)	(1.13)	(0.55)	(-0.00)
R^2	0.904	0.897	0.886	0.867	0.834	0.817
1^{st} stage F-stat	15.26	15.56	15.08	15.72	16.18	12.12
Observations	1400	1400	1400	1400	1400	1400

Notes: The table presents estimates of one measure of innovation (citations received within a five-year window per inhabitants) on the top 1% income share at different lags column (1) uses a one-year lag between the measure of innovation and the dependent variable, column (2) uses two-year lags etc. Both our measure of innovation and the dependent variable are taken in log in all columns. Time span: 1981-2008. Variable description is given in Table 3.

Panel data IV 2SLS regressions with state and year fixed effects. Innovation is instrumented by the number of senators that seat on the appropriation committee. DC is removed from the sample in all columns because it has no senators. The lag between the instrument and the endogeneous variable is set to 3 years. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B8: REGRESSION OF INNOVATION ON TOP 1% INCOME SHARE USING ONLY THE SECOND INSTRUMENT

Dependent variable	Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.206***	0.122***	0.151***	0.167***	0.099***	0.155***
	(3.43)	(3.50)	(3.46)	(3.44)	(3.31)	(2.96)
Gdppc	-0.165*	-0.123	-0.139	-0.117	-0.074	-0.131
	(-1.83)	(-1.41)	(-1.52)	(-1.29)	(-0.84)	(-1.29)
Popgrowth	1.808*	1.351	1.365	1.323	1.297	1.801
	(1.65)	(1.21)	(1.27)	(1.20)	(1.10)	(1.36)
Sharefinance	0.966***	1.259***	1.152***	1.201***	1.449***	1.502***
	(3.86)	(4.99)	(4.75)	(4.87)	(4.69)	(4.03)
Gvtsize	0.555	0.708	0.481	0.540	1.311**	1.944**
	(1.19)	(1.47)	(0.98)	(1.09)	(2.11)	(2.27)
Unemployment	-1.038*	-0.905	-0.804	-0.793	-1.185*	-1.421*
	(-1.71)	(-1.54)	(-1.43)	(-1.41)	(-1.77)	(-1.75)
R^2	0.793	0.831	0.829	0.827	0.806	0.686
1^{st} stage F-stat	81.04	102.62	103.49	93.11	64.98	24.27
Observations	1478	1326	1326	1326	1326	1326

Notes: The table presents estimates of different measures of innovation lagged by two years on the top 1% income share of state income: column (1) uses the number of patents, column (2) uses the number of citations received within a five-year window, column (3) uses the number of claims, column (4) uses the number of patents weighted by their generality index, column (5) uses the number of patents belonging to the top 5% most cited in the year and column (6) uses the number of patents belonging to the top 1% most cited in the year. All these measures as well as the dependent variable are taken in log. Time span: 1983-2011 for columns (1) 1983-2008 for columns (2) to (6). Variable description is given in Table 3.
Panel data IV 2SLS regressions with state and year fixed effects. Innovation is instrumented by a measure of spillover as described in section 6.1. The lags between the instruments and the endogeneous variable is set to 1 year. Control for spatial correlation involves adding two additional controls for demand shocks as explained in subsection 6.1. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B9: Robustness 2: financial sector, natural resources and taxation - IV results

Dependent variable	Top 1% Income Share						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Innovation	0.172**	0.153*	0.173**	0.201*	0.173**	0.174**	0.181**
	(1.99)	(1.92)	(2.02)	(1.74)	(2.05)	(2.02)	(2.03)
Gdppc	-0.211**	-0.199*	-0.215**	-0.219**	-0.199**	-0.214**	-0.232**
	(-2.12)	(-1.91)	(-2.21)	(-2.15)	(-1.97)	(-2.22)	(-2.24)
Popgrowth	-0.009	-0.004	-0.045	0.047	-0.051	-0.022	-0.026
	(-0.01)	(-0.00)	(-0.04)	(0.04)	(-0.04)	(-0.02)	(-0.02)
Sharefinance	0.940**	1.064***	0.943**	1.010**	0.921**	0.945**	0.951**
	(2.38)	(3.31)	(2.27)	(2.07)	(2.07)	(2.27)	(2.27)
Gvtsize	0.115	0.059	0.065	0.069	0.201	0.075	-0.008
	(0.24)	(0.14)	(0.14)	(0.15)	(0.36)	(0.16)	(-0.02)
Unemployment	-1.225*	-1.693**	-1.221*	-1.136*	-1.269*	-1.235*	-1.456*
	(-1.74)	(-2.33)	(-1.79)	(-1.78)	(-1.86)	(-1.79)	(-1.92)
Highways	0.030**	0.021	0.030**	0.032**	0.029**	0.031**	0.031**
	(2.42)	(1.35)	(2.35)	(2.22)	(2.36)	(2.37)	(2.38)
Military	0.011*	0.011*	0.011*	0.012*	0.011*	0.011*	0.010*
	(1.80)	(1.81)	(1.85)	(1.82)	(1.91)	(1.84)	(1.85)
RemunFinance	-0.000						
	(-0.08)						
EFD				-0.660			
				(-1.01)			
Oil					-0.001		
					(-0.14)		
Mining					0.761		
					(1.04)		
Margtax							0.012*
							(1.87)
R^2	0.897	0.907	0.896	0.890	0.897	0.897	0.896
1^{st} stage F-stat	14.04	14.53	14.28	10.56	15.08	14.65	14..43
Observations	1600	1472	1600	1600	1600	1600	1600

Notes: The table presents estimates of the number of citations received with a five year-window per inhabitants lagged by two years on the top 1% income share of state income: in column (1) we control for average compensation in the financial sector, in column (2), NY, CT, DE and MA (the state with the largest financial sectors) are dropped from the dataset, in column (3), finance-related patents have been removed, in column (4) we control for financial dependence in the state as explained in section 6.2, in column (5) we control for the size of oil and mining sectors, in column (6) oil-related patents have been removed in the count of citations and in column (7) we control for the maximum marginal tax rate. Time Span: 1978-2008. Variables *Oil* and *NaturalRessource* measure the share of oil related and natural resources extraction activities in GDP, variable *RemunFinance* measures the compensation per employee in the financial sector, variable *EFD* measures the financial dependence of innovation and variable *MarginalTax* measures the highest marginal tax rate of labor. Time span 1976-2008. Variable description is given in Table 3.
Panel data IV 2SLS regressions with state and year fixed effects. Innovation is instrumented by the number of senators that seat on the appropriation committee. DC is removed from the sample in all columns because it has no senators. The lag between the instrument and the endogeneous variable is set to 3 years. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B10: Robustness 3A: controlling for industry composition - OLS results

Dependent variable	Top 1% Income Share				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Innovation	0.037***	0.029***	0.030***	0.039***	0.026***
	(3.51)	(3.40)	(3.31)	(4.24)	(3.33)
Gdppc	-0.082	-0.092	-0.090	-0.081	-0.089
	(-1.39)	(-1.53)	(-1.50)	(-1.34)	(-1.49)
Popgrowth	-0.121	-0.100	-0.109	-0.363	-0.112
	(-0.12)	(-0.10)	(-0.11)	(-0.37)	(-0.11)
Sharefinance	0.411**	0.426**	0.425**	0.401**	0.424**
	(2.18)	(2.21)	(2.21)	(2.16)	(2.20)
Gvtsize	0.219	0.178	0.169	0.074	0.166
	(0.65)	(0.53)	(0.50)	(0.22)	(0.50)
Unemployment	-0.624	-0.645	-0.624	-0.791*	-0.647
	(-1.43)	(-1.48)	(-1.43)	(-1.82)	(-1.48)
Size of Sector:					
Computer and Electronic				-0.235***	
				(-3.46)	
Chemistry				0.195	
				(1.62)	
Electrical Component				0.535*	
				(1.88)	
R^2	0.915	0.915	0.915	0.917	0.915
Observations	1632	1632	1632	1632	1632

Notes: The table presents estimates of innovation as measured by the number of citations received within a five-year window on the size of the top 1% income share. Innovation is lagged by two years. We look at the effect of industry composition: column (1), excludes patents from the computer sectors (NAICS: 334), column (2) excludes patents from the pharmaceutical sectors (NAICS: 3254) and column (3) excludes patents from the electrical equipment sectors (NAICS: 335), column (4) adds the share of three sectors as additional controls and column (5) excludes citations to patents belonging to three highly exporting sectors: Transportation, Machinery and Electrical Machinery. The size of a sector (see column (4)) is defined as the share of GDP from the corresponding sector. Time span: 1976-2008. Variable description is given in Table 3.
Panel data OLS regressions with state and year fixed effects. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B11: Robustnesess 3B: controlling for industry composition - IV results

Dependent variable	Top 1% Income Share				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Innovation	0.328*	0.174**	0.183**	0.168**	0.160**
	(1.80)	(2.02)	(2.02)	(2.08)	(2.00)
Gdppc	-0.197**	-0.226**	-0.215**	-0.173**	-0.207**
	(-1.97)	(-2.23)	(-2.20)	(-2.09)	(-2.21)
Highways	0.036**	0.030**	0.030**	0.020*	0.029**
	(2.11)	(2.32)	(2.35)	(1.78)	(2.29)
Military	0.018*	0.012*	0.012*	0.009*	0.011*
	(1.92)	(1.86)	(1.88)	(1.71)	(1.80)
Popgrowth	-0.238	0.003	-0.048	-0.494	-0.027
	(-0.20)	(0.00)	(-0.04)	(-0.44)	(-0.02)
Sharefinance	1.129**	0.939**	0.952**	0.729**	0.946**
	(2.03)	(2.27)	(2.26)	(2.22)	(2.26)
Gvtsize	0.716	0.109	0.088	-0.186	0.119
	(0.90)	(0.23)	(0.19)	(-0.43)	(0.25)
Unemployment	-1.413*	-1.268*	-1.157*	-1.267*	-1.307*
	(-1.73)	(-1.80)	(-1.75)	(-1.94)	(-1.82)
Size of Sector:					
Computer and Electronic				-0.501**	
				(-2.48)	
Chemistry				0.240*	
				(1.73)	
Electrical Component				0.052	
				(0.12)	
R^2	0.845	0.890	0.891	0.896	0.890
1^{st} stage F-stat	6.35	14.14	14.93	13.34	14.05
Observations	1548	1548	1548	1548	1548

Notes: The table presents estimates of innovation as measured by the number of citations received within a five-year window on the size of the top 1% income share. Innovation is lagged by two years. We look at the effect of industry composition: column (1), excludes patents from the computer sectors (NAICS: 334), column (2) excludes patents from the pharmaceutical sectors (NAICS: 3254) and column (3) excludes patents from the electrical equipment sectors (NAICS: 335), column (4) adds the share of three sectors as additional controls and column (5) excludes citations to patents belonging to three highly exporting sectors: Transportation, Machinery and Electrical Machinery. The size of a sector (see column (4)) is defined as the share of GDP from the corresponding sector. Time span: 1976-2008. Variable description is given in Table 3.
Panel data IV 2SLS regressions with state and year fixed effects. Innovation is instrumented by the number of senators that seat on the appropriation committee. DC is removed from the sample in all columns because it has no senators. The lag between the instrument and the endogeneous variable is set to 3 years. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B12: Robustness 4: controlling for agglomeration effect - OLS and IV results.

Dependent variable	Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Innovation	0.022**	0.021**	0.020**	0.196*	0.216*	0.232*
	(2.27)	(2.14)	(2.02)	(1.90)	(1.85)	(1.79)
Gdppc	-0.151**	-0.152**	-0.153**	-0.208**	-0.197**	-0.191**
	(-2.34)	(-2.36)	(-2.38)	(-2.48)	(-2.46)	(-2.43)
Popgrowth	-0.300	-0.302	-0.302	-0.241	-0.244	-0.242
	(-0.29)	(-0.30)	(-0.30)	(-0.21)	(-0.21)	(-0.21)
Sharefinance	0.370*	0.371*	0.375*	0.829**	0.828**	0.818**
	(1.85)	(1.86)	(1.87)	(2.14)	(2.11)	(2.08)
Gvtsize	-0.013*	-0.013*	-0.013*	-0.007	-0.005	-0.005
	(-1.89)	(-1.92)	(-1.92)	(-0.81)	(-0.62)	(-0.52)
Unemployment	-0.603	-0.605	-0.611	-1.065*	-1.025*	-0.986*
	(-1.42)	(-1.43)	(-1.44)	(-1.70)	(-1.68)	(-1.65)
Highways				0.031**	0.031**	0.032**
				(2.36)	(2.32)	(2.30)
Military				0.014*	0.016*	0.016*
				(1.88)	(1.86)	(1.83)
Agglo	0.007	0.008	0.009	-0.046	-0.068	-0.084
	(1.27)	(1.17)	(1.22)	(-1.42)	(-1.45)	(-1.44)
R^2	0.916	0.916	0.916	0.895	0.891	0.887
1^{st} stage F-stat				13.97	12.35	10.86
Observations	1632	1632	1632	1600	1600	1600

Notes: The table presents estimates of innovation as measured by the number of citations received within a five-year window on the size of the top 1% income share. Innovation is lagged by two years. We look at the effect of agglomeration as captured by the variable *Agglo*. *Agglo* is the log of the number of firms in the most (columns 1 and 4), the two most (columns 2 and 5), and the three most (columns 3 and 6) innovative sectors for each state and year. Time span: 1976-2008. Variable description is given in Table 3.
Panel data OLS (columns 1 to 3) and IV 2SLS (columns 4 to 6) regressions with state and year fixed effects. DC is removed from the sample in columns (4), (5) and (6) because it has no senators. Innovation is instrumented by the number of senators that seat on the appropriation committee. The lag between the instrument and the endogeneous variable is set to 3 years. t/z statistics in parentheses, computed with autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.