Innovation and Top Income Inequality*

Philippe Aghion Ufuk Akcigit Antonin Bergeaud

Richard Blundell David Hémous

November 2017

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 correlations between measures of innovation and top income inequality. We also show that the correlations between innovation and broad measures of inequality are not significant. Next, using instrumental variable 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 local 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 also thank Daron Acemoglu, Pierre Azoulay, Raj Chetty, Lauren Cohen, 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, John Van Reenen, Fabrizio Zilibotti, 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, the 2016 ASSA meetings and the CEPR-ESSIM 2016 meeting, and finally the referees and editor of this journal for very 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. 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 Magazine, 11 out of 50 are listed as inventors in a US patent and many more manage or own firms that patent, which suggests that these individuals have earned high incomes over time in relation to innovation. 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 looks at the number of granted patents 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.

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.

That the recent evolution of top income inequality should partly relate to innovation, should not come as a surprise. Indeed, if the increase in top income inequality has been pervasive across occupations, it has particularly affected occupations that appear to be closely related to innovation such as entrepreneurs, engineers, scientists but also managers.[2]

We first develop a Schumpeterian growth model where growth results from quality-improving innovations that can be made in each sector either from the incumbent or from a potential entrant. Facilitating innovation or entry increases the entrepreneurial share of income and spurs social mobility through creative destruction as it makes it easier for outsiders to become business owners. In particular, this model predicts that: (i) innovation by entrants and incumbents increases top income inequality; (ii) innovation by entrants increases social mobility; (iii) entry barriers lower the positive effects of entrants' innovations

[1] The worldwide interest for income and wealth inequality, has been spurred by popular books such as Goldin and Katz (2009), Deaton (2013) and Piketty (2014). The sharp increase in top income inequality in the United States over the past decades was documented by Piketty and Saez (2003).

[2] Bakija et al. (2008) find that the income share of the top 1% in the US has increased by 11.2 percentage points between 1979 and 2005, out of this amount, 1.02 percentage points (that is 9.1% of the total increase) accrued to engineers, scientists and entrepreneurs. Yet, innovation also affects the income of managers and CEOs (Balkin et al., 2000; Frydman and Papanikolaou, 2015), and firm owners (Aghion et al., 2016).

on top income inequality and social mobility. Our model also predicts that higher mark-ups for non-innovating incumbents (which may reflect slow diffusion of new technologies or high entry barriers) can lead to higher top income inequality and lower innovation.

We then start our empirical analysis by exploring correlations between innovation and various measures of inequality using OLS regressions. Our main OLS results 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 rate of innovation measured by the flow of citations-weighted patents per capita in this state in this year. Second, innovation is less positively or even negatively correlated with measures of inequality which do not emphasize the very top incomes, or broader measures of inequality like the Gini coefficient, as suggested by Figure 3 below. Next, by looking at the relationship between inequality and innovation at various lags, we find that the correlation between innovation and the top 1% income share is temporary. Finally, the correlation between innovation and top income inequality is dampened in states with higher lobbying intensity.

To make the case that the correlation between innovation and top inequality at least partly reflects a causal effect of innovation on top incomes, we instrument for innovation using data on the Appropriation Committee of the Senate (following Aghion et al., 2009). We argue that the composition of the Appropriation Committee affects the allocation of earmarks across US States, and that this in turn quickly affects patenting and innovation in the states. Then we regress top income inequality on innovation instrumented by the composition of the Appropriation Committee. We find that all the main OLS results in Section 4 are confirmed by the corresponding IV regressions. Our IV results imply that an increase of 1% in the number of patents increases the top 1% income share by 0.2% and that the effects of a 1% increase in the citation-based measures are of comparable magnitude.

Our results pass a number of robustness tests. First, we show that the positive and significant correlation between innovation and top income shares is robust to introducing various proxies reflecting the importance of the financial sector and to controlling for specific sectors' size or for potential agglomeration effects. Second, 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, thereby further validating our instruments.

Next, we use our regression results to calibrate the main parameters of the model, and we use our calibrated model to reproduce the regressions of the paper. We find a very good fit between the OLS and IV regressions coefficients on the one hand, and the coefficients estimated from the calibrated model on the other hand.

Finally, we analyze the relationship between innovation and social mobility using cross-section regressions at the commuting zone (CZ) level. We find that: (i) innovation is positively correlated with upward social mobility (as suggested in Figure 4); (ii) this positive correlation is mainly driven by entrant innovators, and is dampened in CZs 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 by looking explicitly at the effects of innovation on top income shares and social mobility.[3]

Second, it is related to the empirical literature on inequality and growth (see for instance Barro, 2000 who studies the link between overall growth and inequality measured by the Gini coefficient, Forbes, 2000 or Banerjee and Duflo, 2003). More 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. We contribute to this literature by showing that innovation-led growth is a source of top income inequality.

Third, a large literature on skill-biased technical change aims at explaining the increase in labor income inequality since the 1970's.[4] 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. Our results also suggest that innovation does not have a strong impact on broad measures of inequality compared to its impact on top income shares.

Fourth, our paper relates to an active literature on inequality and firm dynamics. Thus 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 their owners and managers. Similarly, Gabaix and Landier (2008) show that the increase in firm size can account for the increase in CEO's pay. 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. Our analysis is consistent with this line of work, to the extent that successful innovation is a main

[3]Relatedly, Hassler and Rodriguez Mora (2000) analyze the relationship between growth and intergenerational mobility in a mulitple equilibria model where growth and social mobility are also positively associated. Yet, in that paper, growth is driven by externalities and not innovations.

[4]Katz and Murphy (1992) and Goldin and Katz (2009) have shown that technical change has been skill-biased in the 20^{th} century. Lloyd-Ellis (1999), Acemoglu (1998, 2002, 2007) or Hémous and Olsen (2016) endogenize the direction of technical change. Krusell et al. (2000) show that in the presence of capital-skill complementarity, the increase in the equipment stock can account for the increase in the skill premium. Several papers (Aghion and Howitt, 1998; Caselli, 1999 and Aghion et al., 2002) argue that General Purpose Technologies (GPT) increase labor income inequality.

factor driving differences in productivities across firms, and therefore in firms' size and pay.[5]

Fifth, our analysis relates to a recent literature on innovation and individuals' income. Thus Frydman and Papanikolaou (2015) find that innovation and executive pay are positively correlated at the firm level. Similarly, Balkin et al. (2000) finds that innovation increases CEO pay in high-tech industries. Aghion et al. (2016) use data from Finland to show that innovation increases an individual innovator's probability to make it to the higher income brackets, and that innovation has an even larger effect on firm owners' income. Bell et al. (2016) find that the most successful innovators see a sharp rise in income. Akcigit et al. (2017) find a positive correlation between patenting intensity and social mobility across US states over the past 150 years.[6]

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.[7] 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. The two papers have in common: (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".[8]

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 state panel data and our measures of inequality and innovation. Section 4 presents our OLS results. Section 5 explains our IV instrument and presents our IV results. Section 6 performs robustness tests. Section 7

[5]Our analysis is also consistent with Hall et al. (2005), Blundell et al. (1999) or Bloom and Van Reenen (2002) who find that innovation has a positive impact on market value.

[6]Gabaix et al. (2016) argue that standard models of individual income dynamics, which are built on random growth, cannot quantitatively account for the increase in income inequality because they generate transitional dynamics that are too slow. Instead, they suggest a model where some agents have a higher mean growth rate. This literature and our paper highlight that such a higher mean could result from innovation.

[7]In Jones and Kim (2014) entrants innovation 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. Moreover, their model predicts a positive correlation between growth and inequality in the short-run (due to a scale effect) and a negative correlation only in the long-run.

[8]Indeed, we show that entrant innovation is positively associated with social mobility. Moreover, while we find that incumbent and entrant innovation contribute to a comparable extent to increasing the top 1% income share, additional regressions in Table B1 of Appendix B suggest that incumbent innovation contributes more to increasing the top 0.1% than entrant innovation (and this holds even more at the top 0.01%).

calibrates the main parameters of the model, and then uses the calibrated model to reproduce our regressions results. Section 8 looks at the relationship between innovation and social mobility. Section 9 concludes. An online appendix with additional theoretical and empirical results and a more detailed description on our calibration exercise and data construction can be found at this link.

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

We consider a discrete time economy populated by a continuum of individuals of measure M. At any point in time a mass $M/(1+L)$ are firm owners and the rest are workers (so that $L \geq 1$ represents the ratio of workers to entrepreneurs). 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:

$$\ln Y_t = \int_0^{\frac{M}{1+L}} \frac{1+L}{M} \ln y_{it} di, \tag{1}$$

where y_{it} is the amount of intermediate input i used for final production at date t. The number of product lines $M/(1+L)$ scales up with population size (as in Howitt, 1999). 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 i at date t, and q_{it} is labor productivity. Each intermediate is produced by a monopolist who faces a competitive fringe.

2.1.2 Innovation

Productive innovation. Whenever there is a new "productive 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. Both entrants and incumbents can undertake productive innovations and we denote by $x_{E,i}$ and $x_{I,P,i}$ their respective productive innovation rates in line i. At the end of period t, other firms can partly imitate the (now 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 from η_H to η_L with $1 < \eta_L < \eta_H$.

Defensive innovation. The incumbent may instead undertake a "defensive innovation" which does not increase productivity (i.e. $q_{i,t} = q_{i,t-1}$) but ensures that she maintains a technological lead of η_H. That is, a defensive innovation prevents the incumbent's potential competitors from using a technology which is too close to hers. We denote by $x_{I,D,i}$ the defensive innovation rate of incumbents. Again, in the absence of a new innovation in period $t+1$, the technological lead of the incumbent shrinks back to η_L.

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.[9]

To innovate with probability $x_{E,i}$ a potential entrant needs to spend

$$C_{E,t}(x) \equiv \frac{\theta_E x_{E,i}^2}{2} \frac{1+L}{M} Y_t;$$

while to undertake productive innovation at rate $x_{I,P,i}$ and defensive innovation at rate $x_{I,D,i}$, an incumbent needs to spend

$$C_{I,t}(x) \equiv \frac{\theta_I (x_{I,P,i} + x_{I,D,i})^2}{2} \frac{1+L}{M} Y_t.$$

The parameters θ_E and θ_I capture R&D productivity for entrants and incumbents respectively, and the innovation cost functions scale up with per capita GDP.

Introducing the dichotomy between *productive* and *defensive* innovations allows us to

[9] 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.

capture the difference between patents and "true innovation": namely, some patents are used to protect rents without contributing much to productivity growth. Indeed, that the observed increase in patenting does not seem to be fully reflected in productivity growth, may be partly due to a growing number of defensive patents.[10]

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).

2.1.3 Timing of events

For simplicity, we rule out the possibility that both entrant and incumbent innovate in the same period.[11] We also assume that in each line i a single potential entrant is drawn from the mass of workers' offsprings. The timing of each period is summarized in Figure 5.

2.2 Solving the model

We solve the model in two steps: in the first step, 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; in the second step, we endogeneize innovation.

2.2.1 Income shares and social mobility for given innovation rates

In this subsection we assume that in all sectors, at any date t, potential entrants innovate at some exogenous rate x_{Et} and incumbents innovate at some exogenous rate x_{It}, knowing that a share ϕ_t of their innovations are productive. Limit pricing in any intermediate sector i implies that the price charged by the incumbent producer is equal to the technological lead η_{it} times the marginal cost $MC_{it} = w_t/q_{i,t}$, hence:

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

[10] An alternative or complementary explanation is that productivity growth from creative destruction may be mismeasured (see Aghion et al., 2017).

[11] Hence, in a given sector, innovations by the incumbent and the entrant are not independent events. This assumption is a discrete time approximation of a continuous time model of innovation. It can be microfounded as follows: Every period there is a mass 1 of ideas, and only one idea is successful. Research efforts x_E and x_I represent the mass of ideas that a firm investigates. Firms can observe each other actions, so that in equilibrium they look for different ideas (as long as θ_E and θ_I are large enough to ensure $x_E^* + x_I^* < 1$).

where $\eta_{i,t} \in \{\eta_H, \eta_L\}$. Innovating allows the technological leader to (temporarily) increase her mark-up from η_L to η_H.

Given the above Cobb-Douglas technology, the final good sector spends the same amount Y_t on all intermediate goods, so that:

$$p_{i,t} y_{it} = \frac{1+L}{M} Y_t \text{ for all } i. \tag{4}$$

Together with (3), this allows us to express equilibrium profits in sector i at time t as

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

Thus equilibrium profits depend only upon mark-ups and aggregate output. Profits are higher whenever the technological leader has recently innovated (no matter the type of innovation, productive of defensive), namely:

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

We can now derive the expressions for the income shares of workers and entrepreneurs. Let μ_t denote the fraction of high-mark-up sectors (i.e. with $\eta_{it} = \eta_H$) at date t. Then, the gross share of income earned by an entrepreneur at time t is equal to:

$$entrepreneur\_share_t = \frac{M}{1+L} \frac{\mu_t \Pi_{H,t} + (1-\mu_t)\,\Pi_{L,t}}{Y_t} = 1 - \frac{\mu_t}{\eta_H} - \frac{1-\mu_t}{\eta_L}. \tag{5}$$

This entrepreneur share is "gross" in the sense that it does not include any potential monetary costs of innovation (and similarly all our share measures are expressed as functions of total output instead of net income—see Appendix A.2 for the expressions of net shares).

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

$$wages\_share_t = \frac{w_t L M}{Y_t\,(1+L)} = \frac{\mu_t}{\eta_H} + \frac{1-\mu_t}{\eta_L}. \tag{6}$$

We restrict attention to the case where $\eta_L - 1 > 1/L$, which ensures that $w_t < \Pi_{L,t}$ for any value of μ_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 a measure of inequality within top-earners).

Since mark-ups are larger in sectors with new technologies, aggregate income shifts from workers to entrepreneurs in relative terms whenever the share of product lines with new technologies μ_t increases. By the law of large numbers this share is equal to the probability of an (unblocked) innovation in any intermediate sector. 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 entry barriers z.

Finally, we measure intergenerational upward social mobility by the probability Ψ_t that the offspring of a worker becomes a business owner. This in turn occurs only if an entrant innovates and is not blocked by the incumbent. Therefore, we get:

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

Social mobility is decreasing in entry barrier intensity z, and it is increasing in the entrant's innovation intensity x_{Et} but less so the higher the entry barrier intensity z.[12] This yields:

Proposition 1 *(i) A higher entrant innovation rate, 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 barrier intensity z; (ii) A higher incumbent innovation rate, x_{It}, is associated with a higher entrepreneur share of income but has no direct impact on social mobility.*

All innovations reduce the wage *share;* however productive innovations increase the equilibrium wage *level* while defensive innovations reduce it.[13] Finally, the entrepreneurial income share is independent of innovation intensities in previous periods, so that a temporary increase in innovation only leads to a temporary increase in the entrepreneurial income share; however, once imitation occurs, the gains from the current burst in innovation will be equally shared by workers and entrepreneurs.

[12] Another interpretation of the decreasing relationship between entry barrier and social mobility is that entry barriers increase the persistence of innovation rents.

[13] Indeed, by plugging (4) and (3) in (1) one gets the equilibrium level of wages as: $w_t = (1+L)\,Q_t/\left(M\eta_H^{\mu_t}\eta_L^{1-\mu_t}\right)$, where Q_t is the quality index defined as $Q_t \equiv \exp\int_0^{\frac{1+L}{M}} \frac{1+L}{M} \ln q_{it} di$. The law of motion for Q_t is given by $Q_t = Q_{t-1}\eta_H^{(\phi x_{It}+x_{Et}(1-z))}$. Therefore, for given technology level at time $t-1$, the equilibrium wage is given by

$$w_t = \frac{1+L}{M} Q_{t-1}\eta_L^{\phi x_{It}+x_{Et}(1-z)-1}\left(\frac{\eta_L}{\eta_H}\right)^{(1-\phi)x_{It}}.$$

This shows that the rate of productive innovations $(\phi x_{It} + x_{Et}\,(1-z))$ increases the contemporaneous level of wage, while the rate of defensive innovations $((1-\phi)\,x_{It})$ decreases it.

2.2.2 Endogenous innovation

We now turn to the endogenous determination of the innovation rates of entrants and incumbents[14]. The offspring of the previous period's incumbent solves the following maximization problem:

$$\max_{x_{I,P},\, x_{I,D}} \left\{ \begin{array}{c} \left(x_{I,P}+x_{I,D}\right)\pi_H + \left(1-x_{I,P}+x_{I,D}-(1-z)\,x_E^*\right)\pi_L \\ +\left(1-z\right)x_E^* \frac{w_t M}{(1+L)Y_t} - \theta_I \frac{\left(x_{I,P}+x_{I,D}\right)^2}{2} \end{array} \right\} \frac{1+L}{M} Y_t.$$

This expression states that the offspring of an incumbent can already collect the profits of the firm she inherited, but also has the opportunity of making higher profit by innovating. Incumbents are indifferent between protective and defensive innovations, so that only the total incumbent innovation rate $x_I = x_{I,P} + x_{I,D}$ is determined in equilibrium (while any share of productive innovation ϕ is an equilibrium).[15] The equilibrium incumbent innovation rate satisfies:

$$x_{I,t} = x_I^* = \frac{\pi_H - \pi_L}{\theta_I} = \left(\frac{1}{\eta_L} - \frac{1}{\eta_H}\right)\frac{1}{\theta_I}, \tag{9}$$

which decreases with the incumbent R&D cost parameter θ_I.

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

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

since a new entrant chooses its innovation rate with the outside option of being a production worker who receives wage w_t. Using equation (6), taking first order conditions, and using our assumption that $w_t < \Pi_{L,t}$ (so that entrant innovate in equilibrium), we can express the entrant innovation rate as

$$x_{E,t} = x_E^* = \left(\pi_H - \frac{1}{L}\left[\frac{\mu_t}{\eta_H} + \frac{1-\mu_t}{\eta_L}\right]\right)\frac{1-z}{\theta_E}. \tag{10}$$

Since in equilibrium $\mu^* = x_I^* + (1-z)\,x_E^*$, the equilibrium entrant innovation rate satisfies:

$$x_E^* = \frac{\left(\pi_H - \frac{1}{L}\frac{1}{\eta_L} + \frac{1}{L}\left(\frac{1}{\eta_L} - \frac{1}{\eta_H}\right)x_I^*\right)(1-z)}{\theta_E - \frac{1}{L}\,(1-z)^2\left(\frac{1}{\eta_L} - \frac{1}{\eta_H}\right)}, \tag{11}$$

[14]Throughout this section, we implicitly assume that θ_I and θ_E are sufficiently large that the aggregate innovation rate satisfies: $x_E^* + x_{I,P}^* + x_{I,D}^* < 1$.

[15]It would be easy to modify the model such that ϕ is uniquely determined: for instance by assuming that $x_{I,P}$ and $x_{I,D}$ are not perfect substitute in the innovation cost function.

so that 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.[16]

Therefore a reduction in either entrants' or incumbents' R&D costs increases innovation and thereby the share of high mark-up sectors and the gross entrepreneurs' share of income. As higher entry barriers dampen the positive correlation between the entrants' innovation rate and the share of high mark-up sectors, 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 entry barriers. We have thus established (proof in Appendix A.1):

Proposition 2 *An increase in incumbent R&D productivity leads to an increase in the incumbent innovation rates x_I^*; an increase in incumbent or entrant R&D productivity leads to an increase in the entrant innovation rates x_E^* but less so the higher entry barriers z; consequently, it leads to higher growth, higher entrepreneur share and higher social mobility but less so the higher entry barriers.*

Here we refer to the entrepreneurial share of income gross of the innovation costs, which amounts to treating those as private utility costs. The results can be extended to the entrepreneurial share net of innovation costs as shown in Appendix A.2.[17]

2.2.3 Extensions

Shared rents from innovation. In the model so far, all the rents from innovation accrue to an individual entrepreneur who fully owns her firm. Yet, our regressions will capture the overall effect of innovation on top income inequality, and in particular the fact that, in the real world, the returns from innovation are shared among several actors (inventors, developers, CEOs, firms' owners, financiers,...). We show this formally in Appendix A.4 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.

[16]The entrant innovation intensity x_E^* increases with x_I^* as 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.

[17]A reason not to include innovation costs is that in practice entrepreneurial incomes are typically generated after these costs are sunk, even though in our model we assume that innovation expenditures and entrepreneurial incomes occur within the same period.

CES production function. We show that our results are robust to the case where (1) is replaced by a CES production function in Appendix A.5.

2.3 From theory to the empirics

2.3.1 Entrepreneurial share and top income share

In our empirical analysis, we shall regress top income shares on innovation. Our innovation measure is based on the number of patents per capita, which is the empirical counterpart of the innovation rate μ_t in the model (the model assumes that the total number of innovations scales up with population size). Our focus so far has been on the entrepreneurial share of income instead of the top income share. Yet, top incomes are earned by entrepreneurs (or more generally individuals associated with innovation) as long as L is sufficiently large. To solve for the top $\alpha\%$ income share, one must consider 3 cases.

Case 1: $\alpha/100 < \mu_t/(1+L)$: The top $\alpha\%$ earners consist only of entrepreneurs who have innovated successfully. Then:

$$Top\_\alpha\%\_share = \frac{\alpha(1+L)}{100}\left(1-\frac{1}{\eta_H}\right).$$

In this case a marginal change in innovation—and therefore μ_t—has no impact the top $\alpha\%$ share.[18]

Case 2: $\mu_t/(1+L) < \alpha/100 < 1/(1+L)$: Then the top $\alpha\%$ earners consist of all entrepreneurs who have innovated successfully plus a fraction of those who have not:

$$Top\_\alpha\%\_share = \mu_t\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right)+\frac{\alpha(1+L)}{100}\left(1-\frac{1}{\eta_L}\right) \tag{12}$$

Thus in this case an increase in the number of (non-blocked) innovations leads to an increase in the top $\alpha\%$ share of income. In particular, we get that:

$$\frac{\partial \ln Top\_\alpha\%\_share}{\partial \ln \mu} = \frac{\mu_t}{Top\_\alpha\%\_share}\left(\frac{1}{\eta_L}-\frac{1}{\eta_H}\right) > 0. \tag{13}$$

If the number of patents per capita is proportional to the number of successful innovations, then the above expression corresponds to the elasticity of the top $\alpha\%$ share with respect to the number of patents per capita. We obtain that, for a given innovation rate, this elasticity

[18] This result depends on our assumption that all innovations have the same size η_H. If one were to relax this assumption and allows for a continuous gap, one would get that an increase in innovation quality would affect the top income share at all percentiles.

is decreasing in α, decreasing in the mark-up of non-innovators η_L and increasing in the mark-up of innovators η_H.

Case 3: $1/(1+L) < \alpha/100$. Then the top $x\%$ earners consist of all entrepreneurs plus some workers. In that case we get:

$$Top\_\alpha\%\_share = \mu_t \left(\frac{1}{\eta_L} - \frac{1}{\eta_H}\right)\left(1 - \left(\frac{\alpha(1+L)}{100} - 1\right)\frac{1}{L}\right) + 1 - \frac{1}{\eta_L} + \left(\frac{\alpha(1+L)}{100} - 1\right)\frac{1}{L}\frac{1}{\eta_L}$$

so that:

$$\frac{\partial \ln Top\_\alpha\%\_share}{\partial \ln \mu} = \frac{\mu}{Top\_\alpha\%\_share}\left(\frac{1}{\eta_L} - \frac{1}{\eta_H}\right)\left(1 - \left(\frac{\alpha(1+L)}{100} - 1\right)\frac{1}{L}\right) > 0. \quad (14)$$

Here as well, an increase in the number of (non-blocked) innovations μ_t leads to an increase in the top $\alpha\%$ share of income, and the corresponding elasticity is increasing in η_H, decreasing in η_L and decreasing in α for given innovation rate.

2.3.2 From inequality to innovation

Although we have emphasized the effect of innovation on top income shares, our 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 (5)) and therefore her incentive to innovate. Thus a higher η_H increases incumbents' (9) and (11) 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, thereby also increasing the entrepreneur share for a given innovation rate. Yet, it decreases incumbents' innovation rate since their net reward from innovation is lower. Under mild conditions (e.g. if $\theta_E \geq (1-z)\,\theta_I/L$), this leads to a decrease in the total innovation rate (see Appendix A.3). Yet, for sufficiently high R&D costs, the overall impact of a higher η_L on the entrepreneur share remains positive. Therefore a higher η_L can contribute to a negative correlation between innovation and the entrepreneur share, leading to a downward bias on the innovation coefficient in an OLS regression of top income inequality on innovation.

2.3.3 Our IV strategy through the lens of our model

Our IV strategy below will rely on shocks which reduce the costs of innovation. In terms of our model, suppose that entrant and incumbent innovation costs are respectively equal to

$\theta_E = \theta\Theta_E$ and $\theta_I = \theta\Theta_I$ where exogenous reductions in θ are driven by our instrument. The causal effect of our instrument on innovation will then be captured by the expression

$$\frac{d\mu_t}{d\theta} = (1-z)\,\frac{dx_E^*}{d\theta} + \frac{dx_I^*}{d\theta}.$$

2.4 Predictions

Overall, the main predictions from the above theoretical discussion can be summarized as follows:

- Innovation by both entrants and incumbents increases top income inequality;
- The effect of innovation on income inequality is stronger on higher income brackets;
- 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, which is in line with findings from the recent literature on the decline of the labor share (e.g. see Elsby et al., 2013 and Karabarbounis and Neiman, 2014).

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 1976, 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 updated Frank-Sommeiller-Price Series from the US State-Level Income Inequality Database (Frank, 2009). From the same data source, we 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) and the Gini Index (definition of these measures can be found in Table 1). These data are available from 1916 to 2013 but we restrict attention to the period after 1976. We end up with a balanced panel of 51 states (we include Alaska, Hawaii and the District of Columbia) over a maximum time period of 36 years. In 2013, the three states with the highest top 1% income share were New-York, Connecticut, and Wyoming with respectively 31.8%, 30.8% and 29.6% whereas Iowa, Hawaii and Alaska were the states with the lowest top 1% income share (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 during the 1970s, with a cross-state coefficient of variation multiplied by 2.2 between 1976 and 2013. Wyoming, Idaho, Montana and South Dakota experienced the fastest growth in the top 1% income share during this time period; while DC, Connecticut, New Jersey and Arkansas experienced the slowest growth.

Income in this database corresponds to 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). It is not possible to decompose total income between its various sources (wage, entrepreneurial or capital incomes) with this dataset, but the World Top Income Database (Alvaredo et al., 2014) gives the composition of the top 1% income share at the federal level. On average between 1976 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 represented 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). 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

A first measure of innovation for each state and each year is the *flow number of patents per capita* in that state and year.[19] For patents granted from 1976, the US patent office

[19] In line with the model, we consider the flow of patents per capita instead of just the flow of patents, to normalize for the size of the state and control for the mechanical fact that larger states innovate more.

(USPTO) provides information on the state of residence of the patent *inventors*, the date of application of the patent and a link to every citing patent. Since a patent may be associated with several co-inventors who may not 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 (yet 85% of the patents we consider are associated with one state only).[20] 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 is above 95%).[21] In addition, we show later that our baseline results are robust to allocating each patent to the state of its assignees (see Appendix B, Table B3).

We associate a patent with its application year which is the year when the provisional application is considered to be complete by the USPTO and a filing date is set. We only consider patents that were ultimately granted by 2014, so that our data suffer from a truncation bias due to the time lag between application and grant. The USPTO estimated in the end of 2012 that patent application data should be considered to be 95% complete for applications filed in 2004.[22] 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). We stop our analysis in 2009 because of the smaller number of patents afterwards.

The annual flow of patent per capita has been multiplied by 1.6 on average between 1976 and 2009. More than 70% of that increase is due to an increase in the number of inventors and 30% is due to an increase in the number of patents per inventor. Yet, simply counting the number of patents granted by their application date is a crude measure of innovation as patents reflect innovations of very heterogeneous quality. However, the USPTO database,

[20]The USPTO classification considers three types of patents: utility patents (for new and useful inventions such as new machines or improvements to existing processes); design patents (for the design of manufactured objects); and plant patents (for new varieties of plants). Utility patents cover 90% of all patents. They are the best proxy for innovation and the only type of patents for which we have complete data, therefore, in line with the literature, we restrict attention to them.

[21]Delaware and DC are the states for which the inventor's address is more likely to differ from the assignee's address for fiscal reasons. See Table B2 in Appendix B for more detail.

[22]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."*

provides sufficiently exhaustive information on patent citations to compute indicators which better measure the quality of innovation. Thus, we consider five additional measures of quality-adjusted innovation rates:

- *Patents per capita weighted by the number of citations within 5 years*: This variable measures the number of citations received within 5 years of the application date. This number has been corrected to account for the different propensity to cite across sectors and time. In addition, because of the drop in the number of observed completed patents in the patent data after 2006, we correct for the truncation bias in citations following Hall et al. (2001). We consider that this series is reliable up to 2006.

- *Patents per capita in the top 5% (or 1%) most cited in a given year.* For each application year, this variable only counts the patents which are among the top 5% (or 1%) most cited patents in the next five years. For the same reasons as above, the corresponding series is stopped in 2006. A rational for using this measure, as argued in Abrams et al. (2013), is the presence of potential non-linearities between the value of a patent and the number of forward citations.

- *Patents per capita weighted by the number of their claims.* The number of claims allows to capture the *breadth* of a patent (see Lerner, 1994, and Akcigit et al., 2016).

- *Patents per capita weighted by their generality.* Following Hall et al. (2001), we compute the generality of a patent as one minus the Herfindahl index of the technological classes that cite the patent, where technological classes are defined at the 4-digit level of the International Patent Classification (IPC). 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 different measures of innovation display consistent trends: thus the four most innovative states between 1975 and 1990 according to the number of patents per capita are also the most innovative states according to the number of (5 years) citations weighted patents per capita, and similarly if we consider the period between 1990 and 2010. From Figure 2, the states which experienced the fastest growth in innovation are Idaho, Washington, Oregon and Vermont; whereas the states with the lowest growth in innovation are West Virginia, Oklahoma, Delaware and Arkansas. More statistics and details are given in Tables 2 to 4.

As pointed out earlier, patenting *per se* may not fully reflect true innovation, but also partly appropriation. Hence the distinction between "productive" and "defensive" innovation in our model above. Moving to more qualitative measures of innovation such as citations, breadth or generality, is meant to also partly address this concern.

3.1.3 Control variables

Regressing top income shares on innovation raises concerns which can be addressed by adding suitable controls. 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, or Bell and Van Reenen, 2014). In turn, the financial sector is sensitive to business cycles and may also affect innovation directly. To address these two concerns, we control for the business cycle via the unemployment rate and for the location specialization index of the financial sector (defined as the share of total GDP accounted for by the financial sector in the state divided by the same share at the national level). 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 can be found in the Bureau of Economic Analysis (BEA) regional accounts and in the Bureau of Labor Statistics (BLS).

Taxation may also create a spurious correlation between top income inequality and innovation as lower taxes could lead to both higher top incomes and higher innovation through the migration of top inventors (see Moretti and Wilson, 2017 for US migration of star inventors and Akcigit et al., 2016 for international migration). To deal with this concern we control for the maximum marginal tax rates on labor and realized capital gains in the state, using data from the NBER TAXSIM project. Agglomeration is also a potential geographical determinant of both innovation and inequality, as we discuss in Section 6.2.

3.2 Estimation strategy

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

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

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 eliminate permanent cross state differences in inequality and aggregate changes.[24] There we are 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 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 1978 and 2011 when measuring innovation by the number of patents, the number of claims or the generality weighted patent count, and from 1978 and 2008 when measuring innovation using the citation based 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 standard errors.[25]

4 Results from OLS regressions

In this section we present the results from OLS regressions of income inequality on innovation. We first look at the correlation between top income inequality and innovation before

[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]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 .758. 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 .812.

[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 B4 in Appendix B.

extending the analysis to other measures of inequality. Next, we see how top income inequality correlates with innovation at different lags. Then, we show how the correlation is affected by the intensity of lobbying, and finally we look at the relationship between inequality and entrant versus incumbent innovation.

4.1 Innovation and top income inequality

Tables 5 and 6 regress the log of the top 1% income share on our measures of innovation. The relevant variables are defined in Table 1. In Table 5, we considered the number of citations per capita in a 5 year window as our measure of innovation and introduce control variable progressively. In Table 6, column 1 uses the number of patents per capita as a measure of innovation, column 2 uses the number of citations per capita in a 5 year window, column 3 uses the number of claims per capita, column 4 uses the generality weighted patent count per capita and columns 5 and 6 use the number of patents among the top 5% and top 1% most cited patents in the year, divided by the state's population. All these values are taken in log and lagged by 2 years.

From these tables we see that the coefficient of innovation is always positive and significant. The coefficient on the citations weighted number of patents is larger than that on the raw number of patents which suggests that particularly the more highly cited patents are associated with the top 1% income share (and these are also the 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.

Moreover, from the coefficient presented in the column 1 of Table 6 and the summary statistics in Table 4, we can compare the magnitude of the correlation between innovation and the top 1% income share with the correlation between the top 1% income share and the importance of the financial sector: thus a one standard deviation increase in our measure of innovation, is associated with a 2.4 point increase in the top 1% income share whereas a one standard deviation increase in the importance of the financial sector is associated with a 1.9 point increase in the top 1% income share. Since as we shall see below, the OLS estimates are likely to be biased, we refer the reader to section 5.1 for further discussion of the magnitude of our effects based on IV regressions.

4.2 Innovation and other measures of inequality

We now run the same regression as before but using broader measures of inequality as dependent variable: the top 10% income share, the Gini coefficient and the Atkinson index. Moreover, with data on the top 1% income share and following Atkinson and Piketty (2007) and Alvaredo (2011), we derive an estimate for the Gini coefficient of the remaining 99% of the income distribution, which we denote by $G99$ where:

$$G99 = (G - top1) / (1 - top1),$$

where G is the global Gini and top1 is the top 1% income share. To check whether the effect of innovation on inequality is 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% share as left hand side variable. This average share is equal to:

$$Avgtop = (top10 - top1) / 9,$$

where top10 represents the top 10% income share.

Table 7 shows the results obtained when regressing these measures of inequalities on innovation. We present results for the citation variable but we get similar results when using other measures of innovation. Column 1 reproduces the results for the top 1% income share. Column 3 uses the top 10% income share, column 2 uses the *Avgtop* measure, column 4 uses the overall Gini coefficient, column 5 uses the Gini coefficient for the bottom 99% of the income distribution and column 6 uses the Atkinson Index with parameter 0.5. We see that innovation: (a) is most significantly positively correlated with the top 1% income share; (b) is less positively correlated with the top 10% income share; (c) is not significantly correlated with the Gini index and is negatively correlated with the bottom 99% Gini. Moreover, the Atkinson index with coefficient equal to 0.5 is positively correlated with innovation.

Finally, in Table 8, we use more concentrated top income share measures, namely the top 0.01, 0.05 and 0.1% income shares. The correlation between innovation and top income share increases as we move up to the income distribution, with the coefficient of innovation reaching 0.087 for the top 0.01% income share.

4.3 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 inventor and assignee disambiguation work of the PatentViews initiative managed by the USPTO[26]. 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.[27] We then aggregate the number of "entrant patents" as well as the number of "incumbent patents" at the state level from 1980.[28]

According to our definition, 17% of patent applications from 1980 to 2014 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). Entrant patents have more citations than incumbent patents: for example in 1980, each entrant patent has 11.4 citations on average whereas an incumbent patent only has 9.5 citations, which is line with the view that entrant patents correspond to more radical innovations (see Akcigit and Kerr, 2010).

Table 9 presents the results from regressing the log of the top 1% income share on 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 both entrant and incumbent innovation are always positive and significant, and the two coefficients are not statistically different from one another.

4.4 Lobbying as a dampening factor

To the extent that lobbying activities help incumbents prevent or delay new entry, we conjecture 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 since 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[29] provides yearly

[26]Accessible online at http://www.patentsview.org

[27]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 B5 in Appendix B). Here, and only here, we focus on patents issued by firms and we have removed patents from public research institutes or independent inventors.

[28]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 to look for the first patent of each assignee, including those with application year before 1976 (but which were granted after 1976).

[29]Data can be found in the OpenSecrets website.

sector-specific lobbying expenditures at the national level from 1998. To proxy for lobbying intensity at the state level, we construct for each state the weighted average of lobbying expenditures in the different sectors (3-digits NAICS sectors), with weights corresponding to sector shares in the state's total employment from the US Census Bureau, a strategy similar to the seminal paper by Bartik (1991).

More precisely, we want to compute the lobbying expenditure in state i in year t, $Lob(i, ., t)$, 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 $emp(i, k, t)$ denotes industry k's share of employment in state i at date t (where $1 \leq k \leq K$ and $1 \leq i \leq I$). From this we compute the aggregate lobbying intensity in state i as:

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

Our measure of lobbying intensity is computed as the logarithm of $Lob(i, ., t)$.

Then we run an OLS regression of the top 1% income share on innovation, this lobbying intensity measure, and the interaction between the two; and we do that separately for entrant innovation (columns 1 to 3 of Table 10) and for incumbent innovation (columns 4 to 6 of Table 10) with the same definition as in previous subsection.

The results are in line with the predictions of our model: we find a negative interaction term between entrant innovation and lobbying intensity. In other words, the effect of entrant innovation on top income inequality is dampened when the lobbying intensity increases.

4.5 Timing between innovation and top income

One may question the choice of two-year lagged innovation in the right-hand side of our baseline regression equation. Here is how we converged on it: first, two years is roughly the average time between a patent application and the date at which the patent is granted by the USPTO as well as most patent offices; second, evidence points at inventors' income moving up immediately after or even before the patent is granted. Thus, using Finnish individual data on patenting and wage income, Toivanen and Väänänen (2012) find an immediate jump in inventors' wages after patent grant. Similar conclusions are obtained respectively by Depalo and Addario (2014) using EPO data and by Bell et al. (2016) using USPTO data: Depalo and Addario (2014) find that inventors' wage peak around the time of the patent application, and Bell et al. (2016) show that the earnings of inventors start increasing even before the filing date of the patent application. In the same vein, Frydman and Papanikolaou (2015) find that executive pay goes up during the year where the patent is granted.

That inventors' income (and more generally innovation-related incomes) should increase even before the patent is granted, is not so surprising after all. First, patent applications are mostly organized and supervised by firms which 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. Second, firms may sell a product embedding an innovation before the patent has been granted, thereby already appropriating some of the profits from the innovation. Similarly, the shareholders of an innovating firm can sell their stocks and thereby benefit from the innovation even before the patent is granted. Third, already at the application stage patenting is associated with easier access to VC financing or with a higher likelihood of an IPO for start-up firms, both of which may translate into a higher income for the innovating entrepreneur (e.g. see Hsu and Ziedonis, 2008 or Haussler et al., 2014).

4.6 Top income inequality and innovation at different time lags

Here we test the robustness of our results to considering to alternative lags for innovation. Table 11 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 2 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 remains significant for up to six-years lags, but with a magnitude that decreases with the lag. The effect eventually disappears as we increase the lag beyond six years. This latter finding is consistent with the view that innovation should have a temporary effect on top income inequality due to imitation and/or creative destruction, in line with the Schumpeterian model in Section 2.[30]

[30]This prediction is likely to be heterogeneous across sectors. For example, the effect is no longer significant after 4 years when restricting to NAICS 336: Transport Equipment, whereas it is still significant after 6 years in sector NAICS 334: Computer and electronic products.

4.7 True innovation or simply appropriation?

The correlations we found so far are between top income inequality and patenting per capita. But patenting per capita is only a proxy for true innovation for at least two reasons. First, a significant proportion of innovations are not patented. Such innovations still induce increases in rents and therefore in top income inequality; yet, to the extent that the benefits from non-patented innovations are less easily appropriable, the relationship between non-patented innovations and top income inequality is likely to be weaker than that between patented innovation and top income inequality. Second, some patents are mostly geared towards preserving incumbents' monopoly rents without contributing much to productivity growth (the "defensive innovations" of our model in Section 2). Yet, two considerations lead us to believe that the correlation we found between patenting and inequality also involves true innovation: (a) defensive innovations are typically made by incumbents, but we showed that entrant innovation is also positively correlated with top income inequality; (b) the correlation between innovations and top income inequality remains strong when we consider more qualitative measures of innovation (number of citations, patent breadth, generality,..), which suggests that it goes beyond a pure appropriation effect of patents.[31]

4.8 Summary

The OLS regressions of top income inequality on innovation performed in this section lead to correlation results that are broadly in line with the predictions of our model, namely: (i) innovation measured by the flow or quality of patenting per capita, 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) top income inequality is positively correlated with both, entrant and incumbent innovation; (v) the correlation between entrant innovation and top income inequality is lower in states with higher lobbying intensity.

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 innova-

[31] In particular, if: (i) changes over time in the share of true innovations among patented innovations remain constant across states; (ii) true patented innovations lead to the same rents as "defensive innovations", then our regressions exactly capture the correlation between top income inequality and true patented innovations.

tion on top income. To reach this conclusion we must account for the possible endogeneity of our innovation measure. Endogeneity could occur in particular through the feedback of inequality to innovation. For example, an increase 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 and main 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. As a robustness test, we will show in Section 6 that this Appropriation Committee instrument can be combined with a second instrument which exploits knowledge spillovers across states.

5.1 Instrumentation using the state composition of Appropriation Committees

We instrument for innovation using the time-varying state composition of Appropriation Committees. To construct this instrument, we gather data on the membership of these committees over the period 1969-2010 (corresponding to Congress numbers 91 to 111).[32]

5.1.1 Some background

The Appropriation Committees of the Senate and of the House of Representatives are standing committees in charge of all discretionary spending legislation through appropriation bills. Discretionary funding are funding that are not required to be allocated to certain program by law (Social Security, unemployment compensation...). This discretionary budget is usually allocated to specific federal departments or agencies. The recipient agency can then disburse these funds to specific projects based on merit and following its own regulations.[33] However, the Appropriation Committees can also choose to add grants (or "earmarks") to the Appropriation Bill to specific projects bypassing the usual peer-review competitive process (e.g. see Aghion et al., 2009; Cohen et al., 2011; Payne, 2003; Savage, 2000; and Feller, 2001).

A member of Congress who sits on an Appropriation Committee often pushes for earmarked grants in the state in which she has been elected, in order to increase her chances of reelection. As a result, state federal research funding to universities in any US State, is

[32]Data have been hand-collected and compared from various documents published by the Senate. The name of each congressman has been compared with official biographical information to determine the appointment date and the termination date of the congressman.

[33]Nevertheless, as mentioned by Payne (2003), a congressman can influence the use of the award by providing funding guidance to the agencies, which they typically comply with.

largely affected by the presence of a congressman from that State in the committee as shown by Payne (2003) and Savage (2000). 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." Evidence that research and research education are large beneficiaries from Appropriation Committees' earmarks, can be found from looking at data from the OpenSecrets project website, which lists the main recipients of the 111$^{\text{th}}$ Congress Earmarks in the US (between 2009 and 2011): universities rank at the top of the recipients list together with defense companies. We shall control for state-level highway and military expenditures in our IV regressions as detailed below.[34]

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

5.1.2 Discussion

We now justify the use of Appropriation Committee Membership as an instrument for innovation. We first argue that the composition of the Appropriation Committee is exogenous. Then, we explain that a nomination to the Appropriation Committee leads to an increase in earmarks received by the state. This boosts innovation in particular because it boosts university patenting which has positive spillovers on innovation in general. We pay particular attention to the timing of each effect.

Exogeneity of the Appropriation Committee Membership A first concern with our instrument is that changes in the state composition of the Appropriation Committee could be related to growth or innovation performance in those states. However, as explained in Aghion et al. (2009), these changes 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. Thus in order to enter the Appropriation Committee, a senator or representative from any state i needs to wait for a seat to become free, and this can happen only if an incumbent is not reelected (or resign, or die) which does not depend on the economic situation in state i.

A related concern is that the composition of the appropriation committee might reflect

[34] See also Aghion et al. (2009), particularly Table 9 and Aghion et al. (2010), Figure 13.

the disproportionate attractiveness - for innovation and wealthy individuals- of states such as California and Massachusetts. However, less advanced states have also 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.[35] Also, the OpenSecrets website shows the cross-state allocation of earmarks from the 111$^{\text{th}}$ Congress: the states that received the highest amount of earmarks per inhabitant, are Hawaii (not too surprising, since the Chairman of the Senate Appropriation Committee at the time, Daniel K. Inouye was a senator from Hawaii) and North Dakota. Other evidence reported by Savage (2000) shows that the top 5 states in terms of academic earmarks in level (not per capita) are Pennsylvania, Oregon, Florida, Massachusetts and Louisiana for fiscal years 1980-1996. The total ranking by earmarks is uncorrelated with the federal research rank and California receives almost the same amount as Hawaii. Cohen et al. (2011) also report a table showing that the states receiving the largest amount of earmarks per capita on average from 1991 to 2008 are HI, AK, WV and MS.[36]

A "zero-stage" regression of earmarks on Appropriation Committee composition To show more systematically how Appropriation Committee Membership affects the allocation of earmarks across US States, we use hand-collected earmarks data gathered from "Citizen Against Government Waste" kindly provided by Cohen et al. (2011). These data associate a state with the "earmark" received during the year by that state. Then we run a "zero-stage regression" of earmarks on Appropriation Committee composition. Formally, we run the following cross-state panel regression:

$$log(E_{i,t}) = \beta_0 + \beta_1 log(E_{i,t-1}) + \beta_2 Senator_{i,t} + X_{i,t}\gamma + B_i + B_t + \varepsilon_{i,t},$$

where t ranges from 1991 to 2008, $E_{i,t}$ denotes the earmarks received by state i during year t per capita, $Senator_{i,t}$ is the number of senators from state i in the Appropriation Committee in year t (thus $Senator_{i,t}$ can only take the values 0, 1 or 2); $X_{i,t}$ are our usual set of covariates; and B_t and B_i are year and state fixed effects.

We run the regression, first using total earmarks as our dependent (LHS) variable and then using only earmarks which we considered to be "research earmarks" based on their

[35] More statistics on the composition of the Senate Appropriation Committee are provided in Table 12.

[36] We tested directly for reverse causality: is it the case that when a state becomes more unequal, it is likely to obtain an additional member in the Appropriation Committee? We run a Probit model with the left hand side variable being a binary variable equal to 1 if a new senator from state i access the committee at t and the LHS variables include the number of senators from state i currently in the committee and the level of the top 1% in log and taken at different lags. We do not find any significant effect of the level of inequality in state i on the probability to access the committee.

title (for example 495,000 dollars have been appropriated to "Energy and Environmental Research Center at the University of ND" in 1991).[37] We present the results in Table 13. We also present similar results using citations received by university patents per capita as the dependent variable, instead of earmarks.[38] Looking at university patents makes sense as earmarks should affect innovation in a state first through their impact on university research. The results are consistent with the existing literature (Payne and Siow, 2003), namely: having one (or two) senator in the committee is associated with increased earmarks and with more and better quality university patents to the corresponding state compared to the US average in the same year.

Timing issue Our IV regression below assumes a three-year lag between the instrument and innovation in the first stage regression. Is this a reasonable assumption?

Consider first the example of Kentucky (KY) with the arrival of the current majority leader (senator McConnell) in the Appropriation Committee in January 1993.[39] Following McConnell's arrival, both earmarks and innovation immediately started to sharply increase. Thus already in 1993 an earmark of more than four million dollars was allocated to the University of Kentucky Advanced Science & Technology Commercialization Center to further develop a business incubator housing new and emerging technology-based companies within the university. And from our earmarks data, we see that the share of total earmarks received by KY underwent a tenfold increase between 1992 and 1993.

McConnell's enrollment in the Appropriation Committee also induced a prompt and substantial increase in patents and citations from that state. To show this, we use a synthetic cohort approach as presented in Abadie et al. (2010). In short, we construct a "synthetic" (or "counterfactual") Kentucky, by pooling a set of other states selected by minimizing the distance in several characteristics between those states and Kentucky before 1993. Figures 6(a) and 6(b) show that the difference in the number of citations to university patents per capita between the actual Kentucky and the "synthetic" Kentucky increases quickly and sharply after Senator McConnell's arrival on the Appropriation Committee in January 1993; while if we consider all patents, the gap between Kentucky and "synthetic" Kentucky widens up three years later.

[37] Our measure of research earmark is a lower bound as there are several recipients which could be research institutions but whose names are not made explicit.

[38] The list of university patents was provided by the USPTO and created by matching the name of the top 250 university with the name of the patent assignee.

[39] Senator McConnell's accession to the committee followed the death of Senator Burdick in 1992. Even if McConnell did not directly replaced Burdick, only four senators were new to the committee when the next congress started operating in January 1993.

Of course this is just one particular example. We then generalize these results by performing an event study exercise, the results of which are reported in Figures 7(a), 7(b) and 7(c). There, we first restrict attention to states that experienced at least one increase in their representation on the Senate Appropriation Committee some time over our sample period.[40] We aggregate the average share of earmarks, citations to university patents and citations to all patents for these states. For any state where such a change occurred "Year 0" refers to the year where the increase in that state's representation on the Appropriation Committee occurred.[41] Figure 7(a) shows that a one-member increase in state representation on the Appropriation Committee translates almost immediately into a sharp increase in the amount of earmarks across states. This is consistent with the findings in Cohen et al. (2011). Figures 7(b) and 7(c) show that university innovation, as measured by a citation-weighted count of patents, also rises quickly after a one-member increase in state representation on the Appropriation Committee and that overall innovation increases three years later.[42]

Finally, our lag choice finds support in the literature. Thus Payne and Siow (2003) find that the appointment of an alumni to the House Appropriation Committee leads to an increase in the number of granted university patents after five years, which corresponds to an increase in patent applications after three years, and we know from Jaffe (1989) that there are large contemporaneous spillovers from university research on corporate patenting. Daim et al. (2007) find a time lag between federal research funding in nanotechnology and patent grants of 5.5 years (which corresponds to a time lag of around 3 years for patent applications); Toole (2007) shows that in the pharmaceutical industry, the positive impact of public R&D on private R&D is the strongest after 1 year; and using shocks to defense R&D, Moretti et al. (2016) show that public R&D expenditures increase private R&D contemporaneously. Finally, Pakes and Schankerman (1984) and Hall et al. (1986) have found very little lag between private R&D and patent applications.

[40] The sample period depends on the measure we consider which in turns affects the number of states in our sample. In addition, in order to increase the size of this sample, we also include states that experienced more than one increase in their representation on the committee, in which case we only consider the first increase event.

[41] Obviously, the actual year of change in Committee representation differs from one state to another, however in this figure we look at the average effect of a one-member increase across all states, abstracting from the heterogeneity in the dates of change.

[42] Those figures also report the mean number of senators around the event—which may differ from 0 pre-event and 1 post-event because here we ignore what happens when senators leave the committee. Moreover, we find that the event has a significant effect on earmarks, in the sense that in Figure 7(a) the sum of the dummies at t+1, t+2, t+3 is significantly different from the dummies at t-1, t-2, t-3 at 5.7 per cent level, the effect on university patents similarly defined is also significant at the 9.2 per cent level, while the effect on patents defined as the sum of the dummies at t+3, t+4, t+5 relative to the sum of the dummies at t-1, t-2, t-3 is significant at 8.3 per cent—these level change to respectively 7.1, 7.3 and 8.8 percent if one adds our set of covariates to the exercise.

Controlling for other expenditures One final concern with our instrument is that not all earmarks go to funding research. For instance, (rich) owners of construction or military companies may capture part of the earmarked funds, given that many earmarks are dedicated to these two sectors. In that case, the number of congressmen seating in the appropriation committee would be correlated with the top 1% income share, but for reasons having little to do with innovation. To deal with this 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 data source, for each state we identify military expenditures and a particular type of infrastructure spending, namely highways, which is presented as a privileged source of earmarks by Aghion et al. (2009). We control for both types of expenditures in our regressions below.

5.2 Regression results

Table 14 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 variable for innovation. In Table 14, 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 (while innovation itself is lagged by 2 years in the main regression) in line with our above discussion. The resulting coefficient on innovation is always positive and significant, and, except for column 6, the F-statistic of the first stage regression is above 10 suggesting that our instrument is reasonably strong.

The results from the first stage regression and the reduced form regression are shown in columns 1 and 2 of Table 15. The coefficient in the reduced form regression suggests that the appointment of an additional senator to the Appropriation Committee increases top income inequality in that state by 1.6%. For the median state-year in terms of GDP (namely State AZ, Year 1990 with a GDP of 103 billion dollars), the top 1% share of fiscal income is 12.5%. Given that roughly half of the GDP ends up as taxable income, we predict a change in income of around 100 million dollars ($0.5 * 103 * 0.016 * 0.125$). As the average yearly earmark in a state with a senator in the Appropriation Committee is equal to roughly 150 million dollars, our regressions results can be accounted for easily without assuming a large multiplier from public R&D to innovation income.[43]

[43]This is all the more true that Delaney (2011) finds that federal earmarks lead to higher state expenditures

5.3 Magnitude

We now consider the magnitude of the impact of innovation on top income inequality implied by Table 14: A 1% increase in the number of patents per capita increases the top 1% income share by 0.22% (column 1 in Table 14) and a 1% increase in the citation-based measures of innovation has a similar effect. This means for example that in California where the flow of patents per capita has been multiplied by 3.2 and the top 1% income share has been multiplied by 2.4 from 1980 to 2005, the increase in innovation can explain 29% 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 23% of the total increase in the top 1% income share over the period 1980-2005.

Yet, one should remain cautious when using our regressions to assess the true magnitude of the impact of innovation on top income inequality. Our coefficient may underestimate the true impact for at least three reasons: (i) the number of citations is a better measure of innovation but is hard to compare over time; (ii) if successful, an innovator from a relatively poor state, is likely to move to a richer state, thereby not contributing 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 different state from the inventors, in which case the effect of innovation on top income inequality will not be fully internalized by the state to which the patent is attributed. Yet, if the share of innovations that get patented is increasing over time, the increase in innovation will be less than the measured increase in patenting, which in turn would mean that the increase in innovation could in fact explain less of the increase in the top 1% income share than what we infer from our regressions.[44]

Looking at cross state differences in a given year, we can compare the effect of innovation with that of other significant variables. Our IV regression suggests that if a state were to move from the first quartile in terms of the number of citations in 2005 to the fourth quartile, its top 1% income share would increase on average by 4.3 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 4.2 percentage-point increase in the top 1% income share.

---

on research education (between 2 and 5 more dollars for each federal dollar).

[44]See the discussion in section 4.7. Although it is a debated topic, 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. Interestingly, our numbers seem consistent with the results of Bakija et al. (2008): they found that 9.1% of the increase in the top 1% income share between 1979 and 2005 accrued to entrepreneurs, technical occupation and scientists, occupations which are the most directly linked to innovation. In addition, innovation increases executives pay and executives, managers, supervisors and business operations (outside the finance sector) account for 38.4% of the increase in the income share of the top 1%.

5.4 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. Yet, the coefficient on innovation is still positive and significant when we only include state and year fixed effects in the regression.[45]

Second, the magnitude of the innovation coefficients in the IV regression is larger than in the OLS regressions. A first potential reason has to do with the relationship between innovation and competition. Our model shows that a higher level of mark-ups for an incumbent who has failed to innovate can 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. More generally, 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. A second reason may have to do with credit constraints: reducing inequality may increase innovation when potential innovators who are not in the top 1% face credit constraints which limit the scope of their innovative investments.[46]

Third, some talented and rich inventors may 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 but through a very different mechanism from the one in our model. However, using disambiguated information on the inventors of patents from the USPTO, we are able to identify the location of successive patents by a same inventor. This in turn allows us to delete patent from inventors that patented in various states. Our results still hold when we look at the effect of patents per capita on the top 1%, with a regression coefficient that is a bit lower than in our baseline (see Table B6 and B7 in Appendix

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

[46]E.g. see Benabou (1996), Aghion and Bolton (1997) and Aghion and Howitt (1998) (Ch. 9). Other mechanisms could explain the gap between the OLS and IV coefficients: for example, a high level of inequality could lead to higher taxes which can harm innovation (Persson and Tabellini, 1994), and inequality can reduce innovation if it reduces the market for innovative goods (Foellmi and Zweimller, 2016).

B respectively for OLS and IV results using only patents by single-state inventors.).

5.5 Other IV results

Appendix B shows 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 B8). Moreover, 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; and it becomes smaller and insignificant from five years (Table B9). These latter findings confirm those in the corresponding OLS Table 11, and indicate again that innovation has a temporary effect on top income inequality.

6 Robustness checks

In this section, we discuss the robustness of our regression results. Robustness concerns may be grouped into three categories. First, potential omitted variables may bias our results. To address this concern, we add more controls in the regression of top income inequality on innovation: namely, the level of financial dependence (subsection 6.1), and a measure of agglomeration (subsection 6.2). This does not alter our results. Second, our results may be driven by one particular economic sector or technology, we successively exclude corresponding sectors or patents to show that this is not the case (subsection 6.1). Finally, one may question the power of our instrumental variable estimations. To address this issue, we add a second instrument based on knowledge spillover and show that the overidentification test does not reject the validity of the two instruments being used jointly (see subsection 6.3).

6.1 Industry composition

Our regressions so far have (mostly) abstracted from industry composition. Here we look more closely at two sectors (Finance and Natural resources) which are likely to have a large impact on top income shares.

The financial sector is heavily represented in the top 1% income share: Bakija et al. (2008) find that 13.2% of primary tax payers belonging to the top 1% worked in the financial sector in 2005 and 22.6% of the increase in the top income share between 1979 and 2005 accrued to individuals in the financial sector. The above regressions already controlled for the share of the financial sector in state GDP, but in Table 17, we perform additional tests in OLS regressions (and Table B11 shows similar results in IV regressions). First, we control for the average employee compensation in the financial sector to capture any direct effect of this variable on the top 1% income share (column 1). Second, we exclude states in which financial activities account for a large fraction of GDP, namely New York, Connecticut, Delaware and Massachusetts (column 2). Third, we exclude financial innovations (patents belonging to the class 705: "Financial, Business Practice") in column 3. 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 (innovation is measured as the number of citations within 5 years per capita). Relatedly financial development may impact both innovation (by providing easier access to credit to potential innovators) and income inequality at the top (by boosting high wages). We build a specific variable to control for this channel. We map patents to 16 NAICS industries and for each state we compute the share of patents in each industry. Then, knowing the industry-level of external financial dependence, we compute the average level of external financial dependence of innovations in each state.[47] This variable (denoted EFD in Table 17) should capture a variation in innovation at state-level driven by a sector that is highly dependent on external finance. Regression results are presented in column 4: the effect of innovation remains significant with a slightly lower coefficient than in the baseline regression. In addition, the effect of the financial dependence of innovations on the top 1% income share is positive and significant.

Natural resources, notably 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 likely to be affected by these sectors which are quite volatile. To address this issue, we control for the share of natural resources and oil extraction in GDP in column 5 of Table 17. Moreover, we remove patents from class 208 (Mineral oils: process and production) and 196 (Mineral oils: Apparatus) in column 6. Here again, our results remain significant.

Finally, we check whether our results are driven by certain sectors which are particularly innovative. We use the mapping between patent technological classes and NAICS sectors to remove patents related to category 334: "Computer and Electronic Products" to exclude the fast-growing computer industries. Similarly, we remove patents from the pharmaceutical sector (NAICS 3254) and from the electrical equipment sector (NAICS 335). Next, we add controls for the share of these three sectors.[48] Then, we use the COMTRADE database to

[47]We use the mapping between technological classes and NAICS codes from the USPTO website. To measure external financial dependence at the industry level, we use the 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 as in Rajan and Zingales, 1998).

[48]In order to obtain complete series, we replace the pharmaceutical sector by the whole chemistry manu-

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 1976 to 2013, we identify three such sectors (Transportation, Machinery and Electrical Machinery), and we check whether our results are robust to excluding them. The results are shown in Appendix B, Table B12, where we conduct OLS regressions using the number of citations within a five-year window to measure innovation. Innovation remains positively and significantly correlated with the top 1% income share, and the coefficient remains stable across specifications.

6.2 Controlling for agglomeration effects

One may wonder whether our results do not reflect agglomeration effects: for example, suppose that some exogenous investment taking place in one particular location (say the Silicon Valley), makes that location more attractive to skilled/talented individuals from other parts of the US. The resulting increased agglomeration of high-skill individuals should lead to both a higher top 1% income share and a higher level of innovation in the corresponding US state, but without the former necessarily resulting from the latter. Glaeser et al. (2009) and Baum-Snow and Pavan (2013) point to higher density cities displaying more income inequality while Carlino et al. (2007) and Gyourko et al. (2013) point at a positive correlation between population density and innovation and the importance of superstar cities.

Figure 2 in the introduction suggests that this should not be such a big concern for our analysis: in particular 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 need a measure of density that is not distorted by large rural areas like upstate New York. In the spirit of Ciccone and Hall (1996), we proceed as follows: in any state, we consider the 10 percent most populated counties in 1970, in 2015, and on average over the period 1970-2015 from the BEA regional accounts. We then compute the population density in these counties every year. This yields three indicators of urban density that are meant to capture the fact that some states were more attractive over the corresponding periods. Running our previous regressions with these additional control variables does not affect our results as seen in Table B13 in the Appendix B (OLS regression results in columns 1, 2, 3 and IV results in columns 4, 5, 6).

facturing sector (NAICS 325).

6.3 Adding a second instrument

To add power to our instrumental variable estimation, 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 lags. Citations reflect past knowledge spillovers (Caballero and Jaffe, 1993), hence a citation network reflects channels whereby future knowledge spillovers occur. Knowledge spillovers in turn lower the costs of innovation (decrease θ_I or θ_E in the model). 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,[49] $w_{i,j,k}$ denotes the relative weight of state j in the citations with lag k of patents issued in state i, aggregated over the period from 1976 to 1978.[50]

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:

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

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 (yet doing so does not impact our results).

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).[51]

[49]In our sample over the period 1976-2014, 67% of backward citations are made to patents filed less than 10 years before the citing patent.

[50]We observe all the patents which received citations from patents granted after 1976 even if the cited patents were granted before 1976 thanks to Hall et al. (2001).

[51]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

Yet, one may worry that this instrument might capture regional or industry trends that are not directly the result of innovation but affect both top income inequality and innovation in that state. For example, a boom in a state may increase local innovation but also lead to more innovation in a neighboring state. Then, if these two states also cite the patents from each other a lot, our spillover variable would capture a positive correlation between innovation in the two states, even though this correlation would mainly reflect a common demand shock. In practice, this concern is mitigated first by the weak correlation between the knowledge spillover weights and geographical distance (below 15%) and second by the (at least) 4 years time lag set between the state innovation measure and the innovation measures used in the instrument. Nevertheless, to proxy for such demand shocks, we build a control variable by computing a weighted average of other states' per capita GDP using an average of the weights calculated before for k between 3 and 10.

Similarly, consider now two states that are highly involved in, say, the computer sector. Then a demand shock in this sector would boost innovation and may increase the top 1% income share in both states, violating our exclusion restriction. The time lag once again mitigates this concern. Yet, to deal with such possibility, we build new weights based on the angular distance between the industry compositions of the manufacturing sectors of the two states. 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 divide this sum by $Pop_{-i,t}$.

A first important finding is that 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%. This in turn reinforces the first instrument.[52]

Table 16 presents the results from the IV regressions of top income inequality on the two instruments combined.[53] As in Table 14, the coefficients are always positive and significant (now at the 1% level). The coefficients are close to those of Table 14, which is all the more remarkable that the two instruments are uncorrelated once one controls for states and time fixed effects. The F-statistic for the two instruments combined, is always above 10.

---

when constructing the weights: the results are essentially unaffected by this change.

[52] 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.

[53] The results from the corresponding first stage and reduced form regressions, are shown in Table 15. In the Appendix, Table B10, we show the results from the IV regressions using only the second instrument.

7 Reproducing our regression results from the model

We now calibrate the main parameters of the model and then use our calibrated model to reproduce the regressions of the paper. Our goal here is two-fold: checking whether our model and our empirical results can be consistent with each other for reasonable parameters, and assessing whether the gap between the OLS and the IV coefficients can be rationalized. We focus on the case where there is no lobbying, i.e. where $z = 0$, so that we are left with 6 parameters to calibrate: the mark-ups η_L and η_H, the R&D parameters θ_I and θ_E, L which is one-to-one related to the share of the population who obtains the monopoly rents (namely $1/(1 + L)$) and the share ϕ of productive innovations among all incumbent innovations (technically ϕ is an equilibrium value but since it is undetermined in equilibrium, we treat as a parameter here). As explained in Section 2.3, we think of the number of innovations in the model as being proportional to the number of patents or citations-weighted patents in the data. We draw three among the six moments directly from the data: the average top 1% share across US states between 1977 and 2011 ($M_1 = 0.13$), the ratio of citations to entrant over incumbent patents ($M_2 = 0.2$), and the elasticity of top income inequality with respect to innovation (we take $M_3 = 0.185$, the coefficient reported in Table 14, column 2).

We then fix the values of three moments from the literature: the average mark-up $M_4 = 1.2$ (according to Jaimovich and Floetotto, 2008, markups range from 1.2 to 1.4 in value added data and from 1.05 to 1.15 in gross output data); the share of employment in entering firms $M_5 = 0.03$ (in line with Garcia-Macia et al., 2016 who find a employment share for entrants of 15% where entrants are defined as firms with less than 5 years) and the growth rate of the economy $M_6 = 0.02$.

The model is fully identified and Appendix C details how each parameter is identified. In a nutshell, the entrant employment share M_5, the semi-elasticity $M_1 * M_3$ and the average mark-up M_4 determine the equilibrium innovation rate μ, the low mark-up η_L and the high mark-up which is also the innovation size η_H. In the relevant case, the semi-elasticity $M_1 * M_3 = \mu \left(\frac{1}{\eta_L} - \frac{1}{\eta_H}\right)$, so that for a given η_L, the semi-elasticity increases both in the innovation rate μ and the innovation step η_H. On the other hand, and for given (harmonic) average mark-ups, the entrant share of employment E increases with the innovation rate μ but decreases with the innovation size η_H. Therefore $M_1 * M_3$ and M_5 together allow to separately identify μ and η_H. The low mark-up η_L is then adjusted to reproduce the average mark-up. Given η_H, η_L and μ, the other parameters are identified through the top 1% share (for L), the innovation ratio and the innovation rate equations (for θ_I and θ_E) and finally the growth rate (for ϕ). Table 18 summarizes the moments that we target, their source, their value in the simulated data described below and gives the value of the different parameters.

The model predicts a large gap between θ_I and θ_E because most innovations are done by incumbents. We find a low ϕ, so that a substantial fraction of incumbents' innovations are "defensive", which is consistent with a large role for innovations in explaining top income inequality, while at the same time measured GDP growth has been timid. With these parameters, the economy is in "case 2" of Section 2, where the top 1% includes all innovators and some incumbent entrepreneurs who failed at innovating. Moreover, with these parameters, an increase in η_L increases the top 1% share but reduces innovation.

We now use our calibrated model to reproduce the regressions of the paper. We consider that there are 51 states over a 28-year time span. In each state i, and in each year t, we assume that the innovation costs for entrants and incumbents are:

$$\theta_{E,i,t} = \theta_E \exp\left(\varepsilon_{\theta,i,t} + \varepsilon_{\theta,i}\right) \text{ and } \theta_{I,s,t} = \theta_I \exp\left(\varepsilon_{\theta,i,t} + \varepsilon_{\theta,i}\right),$$

where the shocks $\varepsilon_{\theta,i,t}$ and $\varepsilon_{\theta,i}$ are respectively state-year and state specific i.i.d shocks. We assume that the markup of non-innovators is given by

$$\eta_{L,i,t} = \eta_L + \varepsilon_{\eta,i,t} + \varepsilon_{\eta,i}$$

where $\varepsilon_{\eta,i,t}$ and $\varepsilon_{\eta,i}$ are respectively state-year and state specific i.i.d shocks. The parameters η_H and L are constant across states and years.

We now compute for each year and state, the innovation rates and the top income shares ($\widehat{Top\_1\%\_share}_{i,t}$) as predicted by our model. We then add "measurement errors" on the top income share, so that

$$Top\_1\%\_share_{i,t} = \widehat{Top\_1\%\_share}_{i,t} \times \exp\left(\varepsilon_{\delta,i} + \varepsilon_{\delta,t} + \varepsilon_{\delta,i,t}\right),$$

where $\varepsilon_{\delta,i}$, $\varepsilon_{\delta,t}$ and $\varepsilon_{\delta,i,t}$ are respectively state, year and state-year specific shocks.

We consider that the number of citations in a state i at a year t is given by

$$Cit5_{i,t} = C\mu_{i,t} \exp\left(\varepsilon_{\mu,i,t}\right)$$

where $\mu_{i,t}$ is the number of innovations, C is a constant and $\varepsilon_{\mu,i,t}$ represent measurement errors. We then run the following regression:

$$\log Top\_1\%\_share_{i,t} = A + B_i + B_t + \beta_1 \log Cit5_{i,t} + \varepsilon_{i,t},$$

first in a simple OLS regression and then in an IV regression where we instrument $Cit5_{i,t}$ by

$\varepsilon_{\theta,i,t} + \varepsilon_{\theta,i}$ (which corresponds to a shock to the innovation technology akin to our appropriation committee instrument).

We set the standard deviations of the different shocks to match second order moments in the data as explained in Appendix C. The OLS and IV coefficients averaged over 500 draws on the simulated data give a coefficient of 0.184 for the IV (close to the target coefficient 0.185 from column 2 of Table 14) and of 0.051 for the OLS, close to the 0.049 figure reported in column 2 of Table 6.[54] Figure 8 in Appendix C plots the whole distribution of the IV coefficients and Table 18 shows the average value for the targeted moments in the simulated data. Therefore overall we get a close mapping between its quantitative predictions and our empirical results.

Finally, this exercise makes it easier to understand the difference between the OLS and the IV coefficients in our regressions. The IV coefficient captures the effect of a shock to innovation costs and therefore the positive impact of innovation on top income inequality. The OLS coefficient captures the overall correlation between innovation and top income inequality, which is less strongly positive if only because the variation in η_L creates a negative relationship between innovation and top income inequality. Moreover, the noise on the citation variable further attenuates the OLS coefficient but that effect is small (without it, the OLS coefficient would be on average equal to 0.070).

8 Innovation and social mobility

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

8.1 From cross-state to CZ-level analysis

State-level 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.

Before we look at social mobility, we check whether the effect of innovation on inequality measures at the CZ level is consistent with our cross-state panel findings. At the CZ level, we do not have direct access to data on top income shares for the whole population. We

[54] The IV coefficient is quite stable as long as the standard errors are not too large. The OLS coefficient depends on how much variation there is in η_L at the state-year level relative to θ.

use the census data from 2000 and 2005-2011 to obtain individual earnings information, but unfortunately the publicly available data are censored for income levels above a threshold. To circumvent this problem, we follow Clemens et al. (2017) and assume a Pareto shape distribution for large incomes, which allows us to *compute* top income shares for 726 CZs (details in Appendix D).

We use the county of the inventor of each patent to associate it with a CZ, we obtained this information from the USPTO from 1998 onwards. Finally, we aggregate county level data on total income, financial and government sectors size, unemployment and population from the BEA and the BLS to compute our control variables.

Regressing top income inequality on innovation at the CZ level allows us to introduce both CZ fixed effects and state$\times$year fixed effects, thereby absorbing any variation in innovation at the state-year level. We add controls for the log of total income per capita, for the growth of total population, for the size of financial and local government sectors compared to the US average and for unemployment. These controls have been selected to match our cross-state analysis as closely as possible given the available data. Standard errors are clustered by state, and CZs are weighted by population to account for potential correlation across neighboring CZs and also to give more weight to urban areas.

We present the results in Table 19, where innovation is measured by the number of patents per capita (we run the regressions over the years 2000 and 2005-2011). We find a positive and significant coefficient for innovation, slightly smaller than in the state-level case.[55]

Yet, there are several limits to this exercise: first, at the CZ level we do not have direct access to top income share data and must instead rely on an estimate based on censored data; second, the time interval for the cross-CZ panel is much shorter than for the cross-state panel;[56] third, since CZs are smaller than states, innovation rents are more likely than before to accrue to individuals who do not reside in the geographical unit where the innovation takes place; fourth, and most importantly, we cannot use the Appropriation Committee instrument at the CZ level since Senators are elected to represent a state, not a CZ.

8.2 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)

[55] There are several CZ with 0 patent and interestingly, the coefficient capturing the extensive margin of innovation (as measured by the index taking the value 1 for CZ with no innovations) is negative and significant, so that CZ with no innovation exhibit less top income inequality.

[56] As a matter of fact, when we measure innovation by the number of citations per capita, the panel is even shorter (only includes years 2000 and 2005-2008), and the coefficient ceases to be significant.

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 around 30 year old whereas the percentile P of parents income is calculated over the period 1996-2000 when the child was around 15 year old. The intensity of innovation in each CZ is measured by the average number of citations per capita, but this time, we take the averages over the period 1998-2008.

A potential concern with these data, is that social mobility is based on the location of the parents not the children, so that the data do not account for children who move to and then innovate in a different location from that of their parents. However, 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 for CZ k. We cluster standard errors by state and weight CZ by population as previously.

Table 20 presents our results for this cross-section OLS regression, using the number of citations as a measure of innovation and the same set of control variables as in the previous subsection to which we add school expenditures per student and the manufacturing employment share of the manufacturing sector (both from Chetty et al., 2014) and the average marginal tax rate. 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 one of the two lowest quintiles. The correlation between innovation and social mobility is more positive and significant for the lowest quintile than for the second lowest one. In fact it becomes insignificant for the third and forth quintiles. 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: the correlation is positive but insignificant at the usual thresholds.

The correlation between innovation and social mobility is economically significant. Column 5 shows that moving from the median CZ to the 75^{th} percentile CZ in innovation intensity (which corresponds to an increase in the number of citations per capita by a factor of 2.5) is associated with an increase of 1.2 percentage points in social mobility at the mean level (namely from 9.6% to 10.8%)—where social mobility is measured by the probability of reaching the top quintile when parents belong to the bottom quintile.

One concern is worth mentioning here: quintiles are defined at the national level and 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 the previous regressions.

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.

8.3 Lobbying, entrant and incumbent innovation

Our empirical results have highlighted the positive association between innovation and social mobility. Yet, our model suggests that the effect of innovation on social mobility should operate mainly through entrant innovation, and that entry barriers should dampen it.

To test these predictions, we conduct the same regression as in the previous section but considering separately entrant and incumbent innovations on the right hand side of the regression equation, where entrants and incumbents are defined as in the cross state case. Table 21 presents our results. Columns 1 to 3 regress our three measures of social mobility on the number of citations received by entrant patents, whereas columns 4 to 6 regress the three measures of social mobility on the number of incumbent citations. 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 and incumbent innovations 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 the same cross-CZ data. We construct lobbying intensity as in the cross state case, building from industry share at the county level. Overall, we are left with 662 CZs which can be separated in two groups of equal size, respectively with high and low lobbying activities. Column 1 (respectively 2) of Table 22 shows the effect of innovation as measured by the number of entrant citations (in log) on the logarithm of absolute upward mobility in CZs above (respectively below) median in terms of lobbying activities. Similarly, columns 3 and 4 include both entrant and incumbent innovation separately. The effect of entrant innovation on social mobility is positive and significant only for CZs that are below median in terms of lobbying intensity. Incumbent innovation has no effect on social mobility, whether we look at CZs 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.

9 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 inequality 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.

Our approach has been to look at the aggregate effect of innovation on top income inequality. This is an essential first step to assess the overall quantitative importance of innovation in top income inequality. Thus our analysis complements more microeconomics studies which explore the relationship between innovation, top income inequality and social mobility using individual data on revenues and patenting.[57]

Our findings also suggest interesting avenues for further research on innovation-led growth, inequality and social mobility. A first extension would be to contrast innovation with other sources of top income inequality, for example from financial and lobbying activities, and look at the effects of these other sources on other measures of inequality and social mobility. Our conjecture is that, unlike innovation, lobbying should be positively correlated with broad measures of inequality, and negatively correlated with social mobility.

Second, our calibration results suggest that our simple model, once enriched to better account for firms' heterogeneity, could be a building block toward a full quantitative model of innovation, firm size distribution and top income inequality. Such a model would be useful to

[57] See Aghion et al. (2016) for such a study using Finnish individual data over the period 1990-2000. See also Toivanen and Väänänen (2012) and Bell et al. (2016) for studies based on individual US data.

assess the contribution of innovation in the rise of market power (De Loecker and Eeckhout, 2017) and also to assess the impact of tax policy, innovation policy (R&D subsidies, patent policy) or competition and entry policy on innovation-led growth and top income inequality. These and other extensions of our analysis in this paper, are left for future research.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller, "Synthetic control methods for comparative case studies: Estimating the effect of Californias tobacco control program," *Journal of the American statistical Association*, 2010, *105* (490), 493–505.

Abrams, David S., Ufuk Akcigit, and Jillian Popadak, "Patent Value and Citations: Creative Destruction or Strategic Disruption?," Working Paper 19647, National Bureau of Economic Research 2013.

Acemoglu, Daron, "Why do new technologies complement skills? Directed technical change and wage inequality," *The Quarterly Journal of Economics*, 1998, *113* (4), 1055–1089.

_ , "Technical change, inequality, and the labor market," *Journal of economic literature*, 2002, *40* (1), 7–72.

_ , "Equilibrium bias of technology," *Econometrica*, 2007, *75* (5), 1371–1409.

_ **, Ufuk Akcigit, and William Kerr**, "The innovation network," Working Paper 22783, National Bureau of Economic Research 2016.

Aghion, Philippe and Patrick Bolton, "A theory of trickle-down growth and development," *The Review of Economic Studies*, 1997, *64* (2), 151–172.

_ **and Peter Howitt**, "A Model of Growth Through Creative Destruction," *Econometrica*, 1992, *60*, 323–351.

_ **and** _ , *Endogenous growth theory*, MIT press, 1998.

_ **, Leah Boustan, Caroline Hoxby, and Jerome Vandenbussche**, "The causal impact of education on economic growth: evidence from US," Technical Report 2009.

_ **, Mathias Dewatripont, Caroline Hoxby, Andreu Mas-Colell, and André Sapir**, "The governance and performance of universities: evidence from Europe and the US," *Economic Policy*, 2010, *25* (61), 7–59.

_ **, Nicholas A. Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt**, "Competition and Innovation: An Inverted-U Relationship," *The Quarterly Journal of Economics*, 2005, *120* (2), 701–728.

_ , **Peter Howitt, and Giovanni L Violante**, "General purpose technology and wage inequality," *Journal of Economic Growth*, 2002, *7* (4), 315–345.

_ , **Timo Boppoart, Peter J. Klenow, and Huiyu Li**, "Missing Growth From Creative Destrution," in "Banque de France - College de France Conference on Secular Stagnation" 2017.

_ , **Ufuk Akcigit, Ari Hyytinen, and Otton Toivanen**, "Living the American dream in Finland: the social mobility of inventors," Technical Report 2016.

Akcigit, Ufuk and William R. Kerr, "Growth Through Heterogeneous Innovations," Working Paper 16443, National Bureau of Economic Research November 2010.

_ , **John Grigsby, and Tom Nicholas**, "The Rise of American Ingenuity: Innovation and Inventors of the Golden Age," Working Paper 23047, National Bureau of Economic Research January 2017.

_ , **Salomé Baslandze, and Stefanie Stantcheva**, "Taxation and the international mobility of inventors," *The American Economic Review*, 2016, *106* (10), 2930–2981.

Alvaredo, Facundo, "A note on the relationship between top income shares and the Gini coefficient," *Economics Letters*, 2011, *110* (3), 274–277.

_ , **Anthony Atkinson, Thomas Piketty, and Emmanuel Saez**, "The world Top Incomes Database," http://wid.world/fr/accueil/ 2014. Accessed: 2016-03.

Armour, Philip, Richard V. Burkhauser, and Jeff Larrimore, "Using The Pareto Distribution To Improve Estimates Of Topcoded Earnings," *Economic Inquiry*, 04 2016, *54* (2), 1263–1273.

Atkinson, A. B. and Thomas Piketty, *Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries* number 9780199286881. In 'OUP Catalogue.', Oxford University Press, 2007.

Bakija, Jon, Adam Cole, and Bradley Heim, "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 March 2008.

Balkin, David B, Gideon D Markman, and Luis R Gomez-Mejia, "Is CEO pay in high-technology firms related to innovation?," *Academy of management Journal*, 2000, pp. 1118–1129.

Banerjee, Abhijit V and Esther Duflo, "Inequality and Growth," *Journal of Economic Growth*, 2003, *8* (3), 267–299.

Barro, Robert J, "Inequality and Growth in a Panel of Countries," *Journal of economic growth*, 2000, *5* (1), 5–32.

Bartik, Timothy J., *Who Benefits from State and Local Economic Development Policies?* number wbsle. In 'Books from Upjohn Press.', W.E. Upjohn Institute for Employment Research, 1991.

Baum-Snow, Nathaniel and Ronni Pavan, "Inequality and City Size," *The Review of Economics and Statistics*, December 2013, *95* (5), 1535–1548.

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

Bell, Brian and John Van Reenen, "Bankers and their Bonuses," *The Economic Journal*, 2014, *124* (574), F1–F21.

Benabou, Roland, "Inequality and growth," *NBER macroEconomics annual*, 1996, *11*, 11–74.

Bloom, Nicholas and John Van Reenen, "Patents, Real Options and Firm Performance," *Economic Journal*, March 2002, *112* (478), 97–116.

Blundell, Richard, Rachel Griffith, and John Van Reenen, "Dynamic Count Data Models of Technological Innovation," *Economic Journal*, March 1995, *105* (429), 333–344.

_ , _ , and _ , "Market share, market value and innovation in a panel of British manufacturing firms," *The Review of Economic Studies*, 1999, *66* (3), 529–554.

Caballero, Ricardo J and Adam B Jaffe, "How high are the giants' shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth," *NBER Macroeconomics Annual*, 1993, *8*, 15–74.

Carlino, Gerald, Satyajit Chatterjee, and Robert Hunt, "Urban density and the rate of invention," *Journal of Urban Economics*, 2007, *61* (3), 389–419.

Caselli, Francesco, "Technological Revolutions," *American Economic Review*, March 1999, *89* (1), 78–102.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of opportunity? The geography of intergenerational mobility in the United States," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Ciccone, Antonio and Robert E Hall, "Productivity and the Density of Economic Activity," *American Economic Review*, March 1996, *86* (1), 54–70.

Clemens, Jeffrey, Joshua D. Gottlieb, David Hémous, and Morten Olsen, "The Spillover Effects of Top Income Inequality," 2017.

Cohen, Lauren, Joshua Coval, and Christopher Malloy, "Do powerful politicians cause corporate downsizing?," *Journal of Political Economy*, 2011, *119* (6), 1015–1060.

Daim, Tugrul, Mitali Monalisa, Pranabesh Dash, and Neil Brown, "Time lag assessment between research funding and output in emerging technologies," *Foresight*, 2007, *9* (4), 33–44.

De Loecker, Jan and Jan Eeckhout, "The Rise of Market Power and the Macroeconomic Implications," Working Paper 23687, National Bureau of Economic Research August 2017.

Deaton, Angus, *The great escape: health, wealth, and the origins of inequality*, Princeton University Press, 2013.

Delaney, Jennifer, "Earmarks and State Appropriations for Higher Education," *Journal of Education Finance*, 2011, *37* (1), 3–23.

Depalo, Domenico and Sabrina Di Addario, "Shedding Light on Inventors' Returns to Patents," Development Working Papers 375, Centro Studi Luca d'Agliano, University of Milano 2014.

Elsby, Michael WL, Bart Hobijn, and Ayşegül Şahin, "The decline of the US labor share," *Brookings Papers on Economic Activity*, 2013, *2013* (2), 1–63.

Feller, Irwin, "Elite and/or Distributed Science," *Innovation Policy in the Knowledge-Based Economy*, 2001, *23*, 189.

Foellmi, Reto and Josef Zweimller, "Is Inequality Harmful for Innovation and Growth? Price versus Market Size Effects," Economics Working Paper Series, University of St. Gallen, School of Economics and Political Science July 2016.

Forbes, Kristin J., "A Reassessment of the Relationship between Inequality and Growth," *American Economic Review*, September 2000, *90* (4), 869–887.

Frank, Mark W, "Inequality and growth in the United States: Evidence from a new state-level panel of income inequality measures," *Economic Inquiry*, 2009, *47* (1), 55–68.

Frydman, Carola and Dimitris Papanikolaou, "In Search of Ideas: Technological Innovation and Executive Pay Inequality," Working Paper 21795, National Bureau of Economic Research December 2015.

Gabaix, Xavier and Augustin Landier, "Why has CEO pay increased so much?," *The Quarterly Journal of Economics*, 2008, *123* (1), 49–100.

_ , Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "The dynamics of inequality," *Econometrica*, 2016, *84* (6), 2071–2111.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J. Klenow, "How Destructive is Innovation?," Working Paper 22953, National Bureau of Economic Research December 2016.

Glaeser, Edward L., Matt Resseger, and Kristina Tobio, "INEQUALITY IN CITIES," *Journal of Regional Science*, 2009, *49* (4), 617–646.

Goldin, Claudia Dale and Lawrence F Katz, *The race between education and technology*, Harvard University Press, 2009.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai, "Superstar Cities," *American Economic Journal: Economic Policy*, November 2013, *5* (4), 167–99.

Hall, Bronwyn H, Adam Jaffe, and Manuel Trajtenberg, "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools," Working Paper 8498, National Bureau of Economic Research October 2001.

_ , _ , and _ , "Market value and patent citations," *RAND Journal of Economics*, 2005, pp. 16–38.

_ , Zvi Griliches, and Jerry Hausman, "Patents and R&D: Is There a Lag?," *International Economic Review*, 1986, *27* (2), 265–83.

Hassler, John and Jose V. Rodriguez Mora, "Intelligence, Social Mobility, and Growth," *American Economic Review*, September 2000, *90* (4), 888–908.

Haussler, Carolin, Dietmar Harhoff, and Elisabeth Mueller, "How patenting informs VC investors The case of biotechnology," *Research Policy*, 2014, *43* (8), 1286–1298.

Hémous, David and Morten Olsen, "The Rise of the Machines: Automation, Horizontal Innovation and Income Inequality," Working Paper 10244, CEPR 2016.

Howitt, Peter, "Steady Endogenous Growth with Population and R & D Inputs Growing," *Journal of Political Economy*, August 1999, *107* (4), 715–730.

Hsu, David H and Rosemarie H Ziedonis, "Patents as Quality Signals for Entrepreneurial ventures," in "Academy of Management Proceedings," Vol. 2008 Academy of Management 2008, pp. 1–6.

Jaffe, Adam, "Real Effects of Academic Research," *American Economic Review*, 1989, *79* (5), 957–70.

Jaimovich, Nir and Max Floetotto, "Firm dynamics, markup variations, and the business cycle," *Journal of Monetary Economics*, 2008, *55* (7), 1238–1252.

Jones, Charles I. and Jihee Kim, "A Schumpeterian Model of Top Income Inequality," Working Paper 20637, National Bureau of Economic Research October 2014.

Karabarbounis, Loukas and Brent Neiman, "The global decline of the labor share," *The Quarterly Journal of Economics*, 2014, *129* (1), 61–103.

Katz, Lawrence F and Kevin M Murphy, "Changes in relative wages, 1963–1987: supply and demand factors," *The Quarterly Journal of Economics*, 1992, *107* (1), 35–78.

Kneer, Christiane, "A Schumpeterian Model of Top Income Inequality," Working Paper 391, Netherlands Central Bank Research Department 2013.

Kortum, Samuel and Josh Lerner, "What is behind the recent surge in patenting?," *Research policy*, 1999, *28* (1), 1–22.

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante, "Capital-skill complementarity and inequality: A macroeconomic analysis," *Econometrica*, 2000, *68* (5), 1029–1053.

Lerner, Josh, "The Importance of Patent Scope: An Empirical Analysis," *RAND Journal of Economics*, 1994, *25* (2), 319–333.

Lloyd-Ellis, Huw, "Endogenous Technological Change and Wage Inequality," *American Economic Review*, March 1999, *89* (1), 47–77.

Moretti, Enrico and Daniel J. Wilson, "The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists," *American Economic Review*, July 2017, *107* (7), 1858–1903.

_ , Claudia Steinwender, and John Van Reenen, "The intellectual spoils of war? Defense R&D, productivity and spillovers," Working Paper, London School of Economics 2016.

Pakes, Ariel and Mark Schankerman, "The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources," in "R&D, Patents, and Productivity," National Bureau of Economic Research, Inc, 1984, pp. 73–88.

Payne, A Abigail, "The effects of congressional appropriation committee membership on the distribution of federal research funding to universities," *Economic Inquiry*, 2003, *41* (2), 325–345.

Payne, Abigail A. and Aloysius Siow, "Does Federal Research Funding Increase University Research Output?," *The B.E. Journal of Economic Analysis & Policy*, May 2003, *3* (1), 1–24.

Persson, Torsten and Guido Enrico Tabellini, "Is Inequality Harmful for Growth?," *The American Economic Review*, 1994, *84* (3), 600–621.

Philippon, Thomas and Ariell Reshef, "Wages and human capital in the US finance industry: 1909–2006," *The Quarterly Journal of Economics*, 2012, *127* (4), 1551–1609.

Piketty, Thomas, *Capital in the twenty-first century*, Harvard University Press, 2014.

_ **and Emmanuel Saez**, "Income inequality in the United States, 1913–1998," *The Quarterly Journal of Economics*, 2003, *118* (1), 1–41.

Rajan, Raghuram G and Luigi Zingales, "Financial Dependence and Growth," *American Economic Review*, June 1998, *88* (3), 559–586.

Romer, Paul M, "Endogenous technological change," *Journal of political Economy*, 1990, *98* (5, Part 2), S71–S102.

Rosen, Sherwin, "The Economics of Superstars," *The American Economic Review*, 1981, *71* (5), 845–858.

Savage, James D, *Funding science in America: Congress, universities, and the politics of the academic pork barrel*, Cambridge University Press, 2000.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter, "Firming Up Inequality," Working Paper 21199, National Bureau of Economic Research May 2015.

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

Toole, Andrew A, "Does public scientific research complement private investment in research and development in the pharmaceutical industry?," *The Journal of Law and Economics*, 2007, *50* (1), 81–104.

Tables and Figures for
"Innovation and Top Income Inequality"

Figures

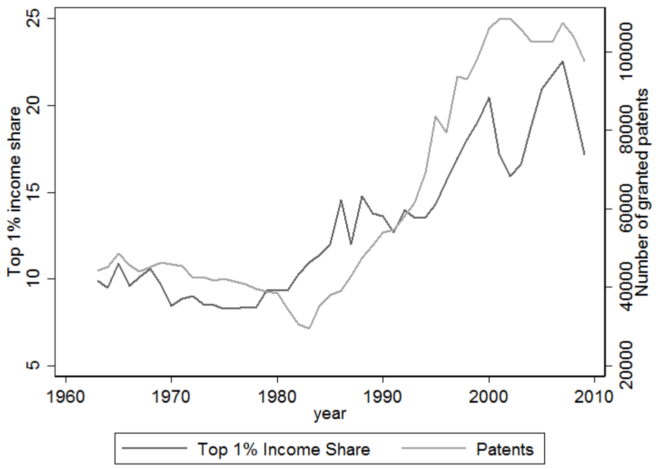


Figure 1: Innovation and Top 1% income share in the US. 1963-2010

Notes: This figure plots the number of granted patent distributed by their year of application against the top 1% income share for the USA as a whole. Observations span the years 1963-2009. Top 1% income shares come from Frank (2009) and patent data come from the USPTO.

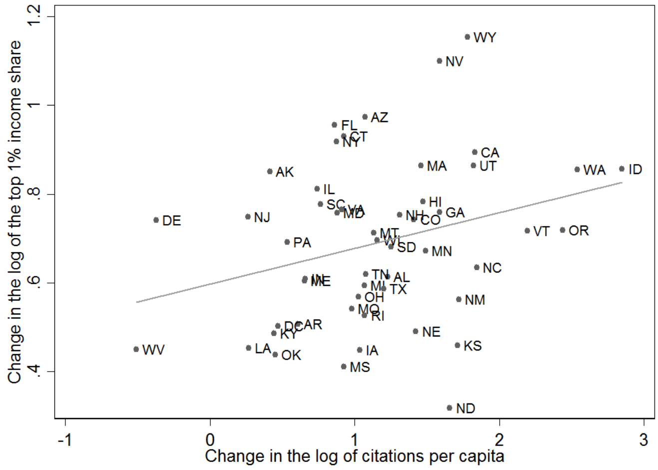


Figure 2: Evolution of innovation and inequality 1980-2005

Notes: 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.

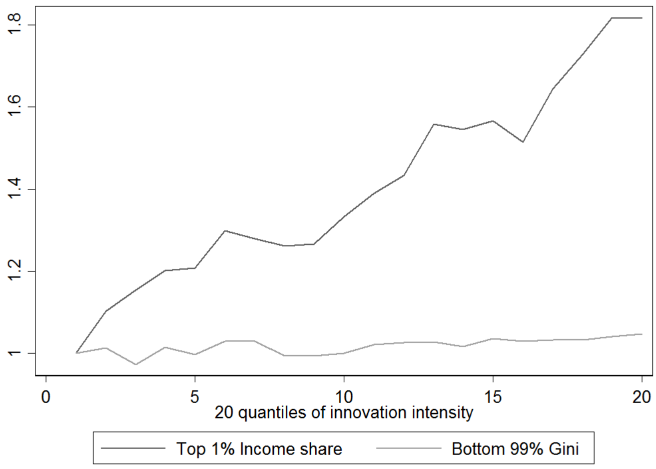


Figure 3: Top 1% income share and Gini coefficient against innovation

Notes: This figure plots the average top-1% income share and the bottom 99% Gini index as a function of their corresponding innovation quantile measured from the number of citations per capita. The bottom 99% Gini is the Gini coefficient when the top 1% of the income distribution is removed. Innovation quantiles are computed using the US state-year pairs from 1976 to 2009. Each series is normalized by its value in the lowest innovation quantiles.

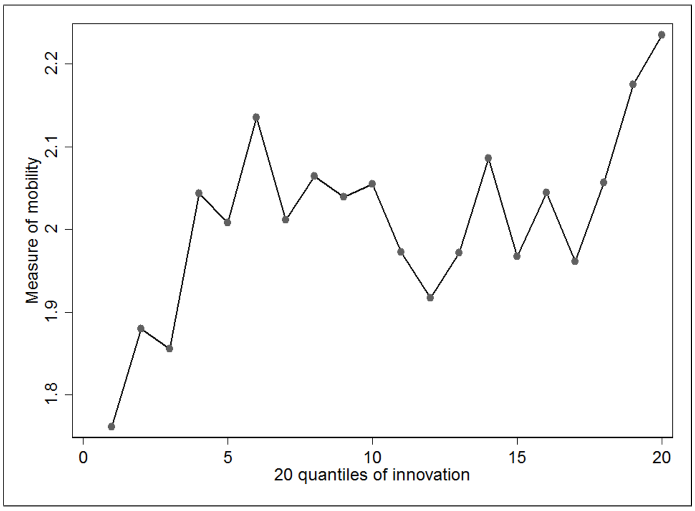


Figure 4: Innovation and social mobility

Notes: This figures plots the percentile in the number of patent per capita (x-axis) against the level 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) and is taken in log. Observations are computed at the Commuting Zones level (677 observations). The number of patents is averaged from 2005 to 2009.

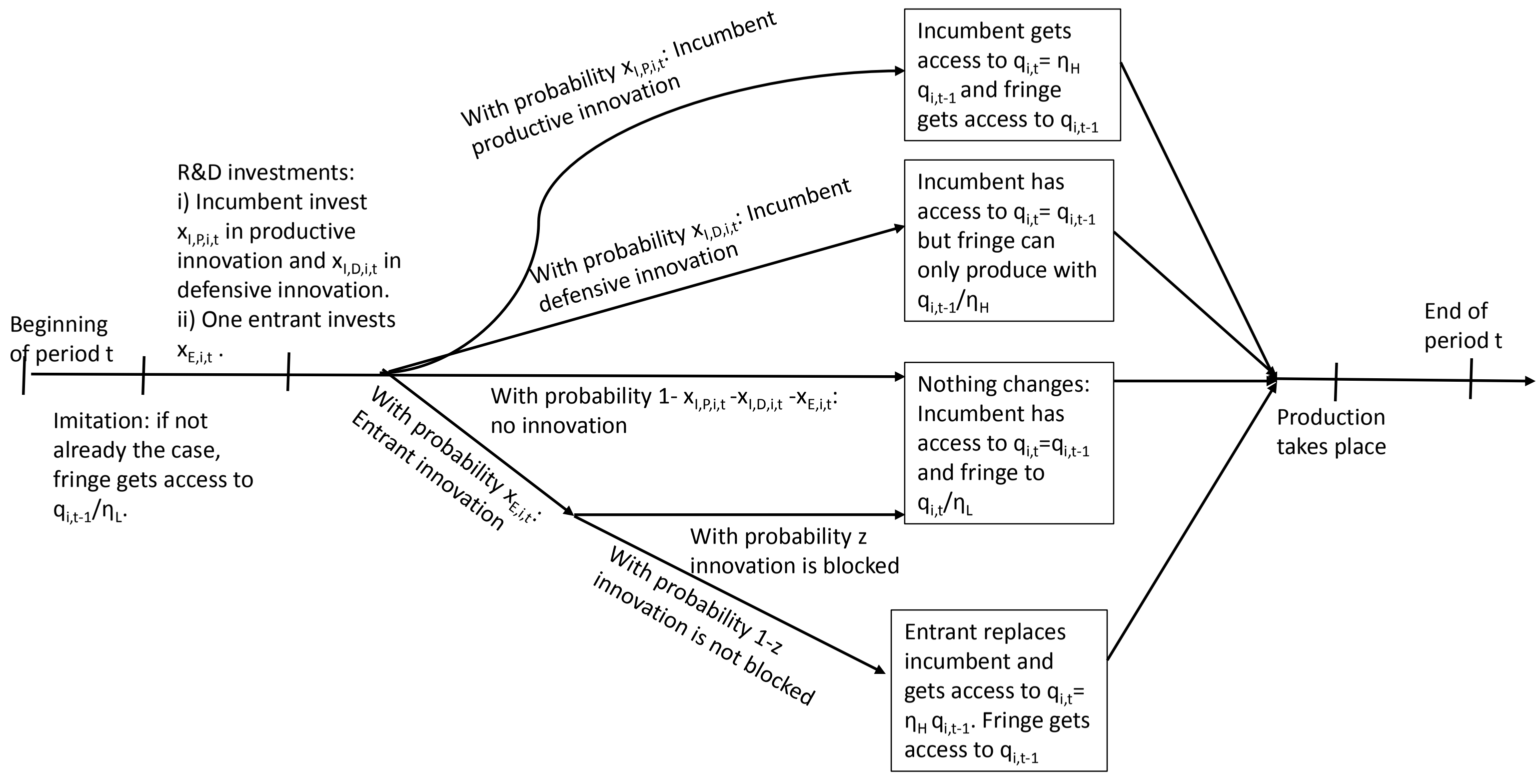
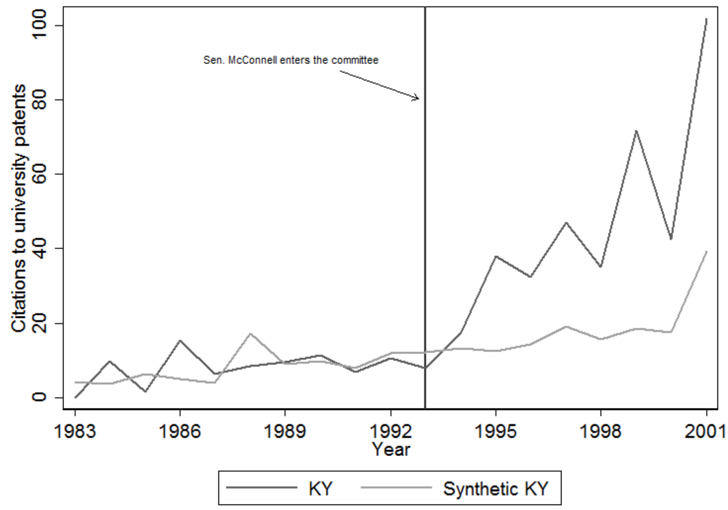


Figure 5: Timeline of events in theoretical model

Notes: This figure shows the timing of events as described in the theoretical model in Section 2.

(a) University patents



(b) All citations

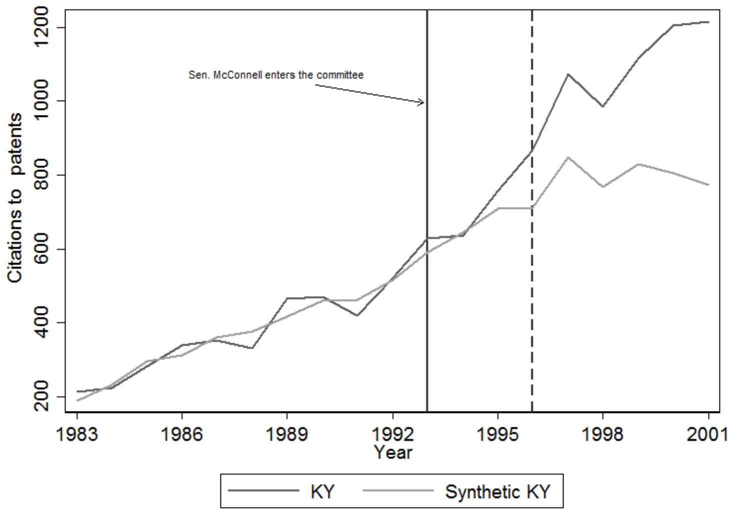
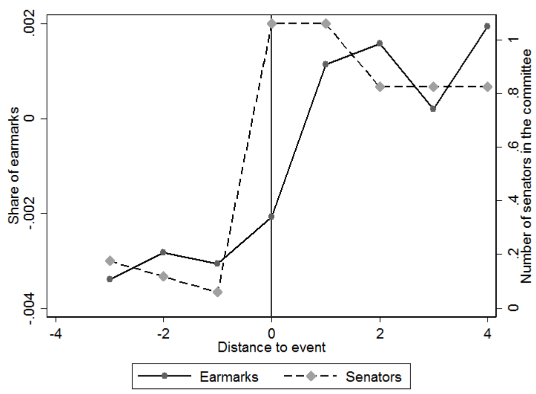


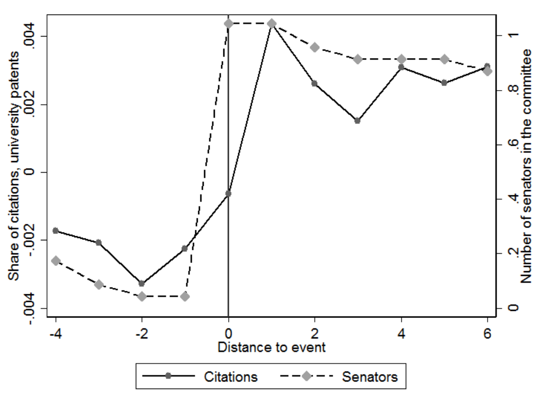
Figure 6: Synthetic cohort analysis

Notes: This figure plots the number of citations received within a 5 year window by all patents (left-hand side figure) and restricting to university patents (right-hand side figure) for the Kentucky and a synthetic Kentucky built by minimizing the distance in terms of size of financial sector, size of public sector, size of the manufacturing sector, GDP growth rate and user cost of R&D taken from Moretti and Wilson (2017). Minimization has been conducted from 1983 to 1991. Treatment year, corresponding to the arrival of senator McConnell in the appropriation committee is 1993. The list of university patents has been received directly from the USPTO.

(a) Earmarks



(b) Citations to university patents



(c) Citations to all patents

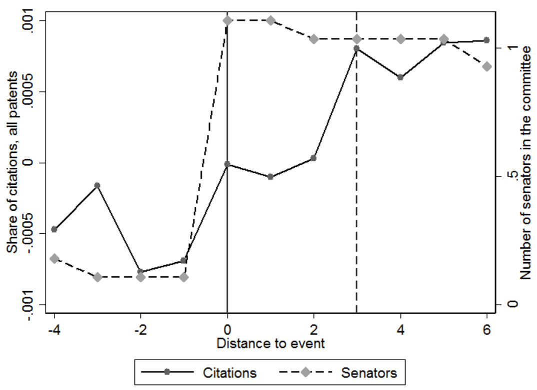


Figure 7: Event studies

Notes: This figure plots the average share of earmarks, citations to university patents and all citations at different time around the appointment of a senator in the appropriation committee. All measures have been residualized on a state specific time trend. Sample is restricted to years 1991-2008 for the top-left panel, 1980-2006 for the top-right panel and 1976-2006 for the bottom panel and to states which experienced one and only one positive change in their representation in the committee. The vertical solid line corresponds to the arrival of a new senator in the committee, the dash line corresponds to 3 years after this event. The list of university patents has been received directly from the USPTO. The list of earmarks has been received from Cohen et al. (2011).

Tables

Table 1: Variable description and notation

Variable names	Description
	Measures of inequality
Top i%	Share of income own by the top i% (i being equal to 1, 5, 10, 0.01, 0.1...) of the income distribution.
Avgtop	Average income share for the percentiles 10 to 2 in the income distribution.
Gini	Gini index of inequality. The Gini index measures the dispersion of the income distribution.
G99	Gini index restricted to the bottom 99% of income distribution.
Atkinson	Atkinson index of inequality with an inequality aversion parameter of 0.5. The Atkinson index is a measure of the gain in terms of utility that would be gained if a total redistribution of the income distribution were to be done.
	Measures of innovation
Patent	Number of patents granted by the USPTO per capita.
Cit5	Total number of citation received no longer than 5 years after applications per capita.
Claims	Total number of claims associated with patents per capita.
Generality	Total number of patents weighted by the generality index per capita.
Top5	Number of patents in the top 5% most cited per capita.
Top1	Number of patents in the top 1% most cited per capita.
	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 GDP accounted for by the financial sector divided by the same share at the country level.
Unemployment	Unemployment rate. Between 0 and 1.
Gvtsize	Share of GDP accounted for by the government sector divided by the same share at the country level.
TaxK	State maximal marginal tax rate for realized capital gains.
TaxL	State maximal marginal tax rate for labor income.
	Additional control variables at the CZ level
Tax	Total tax revenue per capita divided by mean household income per capita for working age adults.
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 2: Descriptive statistics by state in two distinctive years

	1980		2005			1980		2005	
	Top 1%	**Innovation**	**Top 1%**	**Innovation**		**Top 1%**	**Innovation**	**Top 1%**	**Innovation**
AK	5.33	29	12.47	44	**MT**	8.02	46	16.35	144
AL	10.01	40	18.48	138	**NC**	9.03	68	17.04	428
AR	10.05	37	16.68	69	**ND**	9.62	58	13.23	303
AZ	8.56	174	22.66	511	**NE**	9.33	46	15.24	192
CA	9.91	252	24.20	1571	**NH**	8.48	229	18.01	853
CO	9.31	209	19.56	855	**NJ**	9.83	475	20.77	618
CT	12.24	417	31.02	1051	**NM**	8.90	55	15.63	310
DC	14.48	100	23.94	160	**NV**	11.09	118	33.30	574
DE	10.19	588	21.38	406	**NY**	12.08	229	30.25	549
FL	12.23	104	31.78	246	**OH**	8.98	208	15.86	582
GA	8.95	64	19.11	310	**OK**	11.44	228	17.74	357
HI	7.52	29	16.47	125	**OR**	8.25	109	16.91	1251
IA	8.24	113	12.92	318	**PA**	9.37	218	18.71	373
ID	7.68	86	18.08	1483	**RI**	10.25	133	17.36	389
IL	9.63	220	21.67	462	**SC**	8.16	76	17.74	162
IN	8.44	179	15.52	346	**SD**	8.58	32	16.94	113
KS	10.17	74	16.09	410	**TN**	10.09	75	18.76	219
KY	9.69	83	15.76	129	**TX**	12.18	169	21.90	562
LA	11.22	63	17.65	82	**UT**	7.79	124	18.49	768
MA	10.03	324	23.79	1392	**VA**	7.97	119	17.12	295
MD	8.13	183	17.34	441	**VT**	7.97	180	16.31	1614
ME	8.55	73	15.66	140	**WA**	8.37	134	19.69	1689
MI	8.91	233	16.12	678	**WI**	8.21	167	16.48	529
MN	9.31	260	18.24	1154	**WV**	9.54	81	14.97	49
MO	9.96	85	17.11	228	**WY**	9.00	27	28.52	161
MS	10.48	21	15.81	52					

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

Table 3: Descriptive statistics by measures of innovation and for the 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	140	71	113	186	27	501
Cit5	146	64	113	209	21	588
Claims	1347	655	1,039	2,005	222	5,423
Generality	27	12	20	36	3	132
Top5	8	3	5	12	0	41
Top1	3	1	2	4	0	13
2005	Mean	p25	p50	p75	Min	Max
Top 1%	19.07	16.12	17.65	20.77	12.47	33.30
Patents	296	131	230	403	47	904
Cit5	508	161	373	618	44	1,89
Claims	4,67	1,14	3,45	5,06	630	24,64
Generality	104	50	82	152	19	366
Top5	10	2	7	12	0	36
Top1	2	1	1	3	0	9

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 1). All innovation measures are taken per million of inhabitants.

Table 4: MEAN AND STANDARD DEVIATION OF THE MAIN VARIABLES

	Mean	Standard Deviation
Top 1% (log)	2.591	0.267
Patent (log)	5.105	0.764
Cit5 (log)	5.678	1.074
Gdppc (log)	10.49	0.319
Popgrowth	0.010	0.011
Finance	0.920	0.237
Government	1.033	0.273
Unemployment	5.940	2.051
TaxK	4.386	2.948
TaxL	5.297	3.267

Notes: Mean value and standard deviation for the main variables calculated over the period 1980-2005 (relevant variables are defined in Table 1). GDP per capita is calculated in $ per capita and the innovation measures are taken per million of inhabitants.

Table 5: Top 1% income share and innovation

Dependent variable	Log of Top 1% Income Share				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation	0.039***	0.036***	0.043***	0.045***	0.049***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Gdppc		0.075	0.055	0.052	0.063
		(0.054)	(0.053)	(0.052)	(0.044)
Popgrowth		1.146	1.255*	0.864	1.089
		(0.720)	(0.731)	(0.739)	(0.700)
Finance			0.110***	0.118***	0.109***
			(0.041)	(0.043)	(0.036)
Government			-0.007	-0.010	-0.019*
			(0.013)	(0.013)	(0.011)
Unemployment				-0.006*	-0.006*
				(0.004)	(0.003)
TaxK					-0.039***
					(0.004)
TaxL					0.014**
					(0.006)
R^2	0.880	0.882	0.884	0.884	0.896
Observations	1581	1581	1581	1581	1581

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by two years. The dependent variable is the log of the top 1% income share. Panel data OLS regressions with state and year fixed effects. Time span for innovation: 1976-2006. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 6: Top 1% income share and innovation

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.031***	0.049***	0.018**	0.023**	0.024***	0.019***
	(0.011)	(0.009)	(0.009)	(0.010)	(0.005)	(0.004)
Gdppc	0.089**	0.063	0.095**	0.094**	0.076*	0.087**
	(0.043)	(0.044)	(0.045)	(0.043)	(0.043)	(0.043)
Popgrowth	0.943	1.089	0.944	0.916	0.990	1.072
	(0.654)	(0.700)	(0.652)	(0.649)	(0.689)	(0.686)
Finance	0.080**	0.109***	0.072**	0.077**	0.096***	0.091***
	(0.035)	(0.036)	(0.035)	(0.035)	(0.035)	(0.035)
Government	-0.018	-0.019*	-0.017	-0.018	-0.018	-0.016
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Unemployment	-0.006**	-0.006*	-0.005*	-0.006*	-0.006*	-0.005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
TaxK	-0.038***	-0.039***	-0.038***	-0.038***	-0.038***	-0.037***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
TaxL	0.017***	0.014**	0.017***	0.017***	0.013**	0.012**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
R^2	0.889	0.896	0.889	0.889	0.895	0.895
Observations	1734	1581	1734	1734	1581	1581

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by two years. The dependent variable is the log of the top 1% income share. Panel data OLS regressions with state and year fixed effects. Time span for innovation: 1976-2009 (columns 1, 3 and 4) and 1976-2006 (columns 2, 5 and 6). Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 7: Innovation and various measures of inequality

Dependent variable	Top 1%	Top 10 %	Avgtop	Overall Gini	G99	Atkinson
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation	0.049***	0.022***	0.007	-0.001	-0.010**	0.017***
	(0.009)	(0.006)	(0.005)	(0.003)	(0.005)	(0.004)
Gdppc	0.063	0.032	0.002	0.004	-0.021	0.131***
	(0.044)	(0.028)	(0.030)	(0.024)	(0.028)	(0.029)
Popgrowth	1.089	0.553	0.265	-0.382**	-0.553**	0.402
	(0.700)	(0.424)	(0.381)	(0.184)	(0.240)	(0.276)
Finance	0.109***	0.066***	0.021	0.011	-0.018	0.037**
	(0.036)	(0.020)	(0.017)	(0.012)	(0.015)	(0.018)
Government	-0.019*	-0.005	0.013*	-0.004	0.001	-0.029***
	(0.011)	(0.007)	(0.007)	(0.004)	(0.005)	(0.006)
Unemployment	-0.006*	-0.001	0.002	-0.000	0.002	-0.001
	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)	(0.001)
TaxK	-0.039***	-0.018***	-0.002	-0.007***	-0.001	-0.018***
	(0.004)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)
TaxL	0.014**	0.007*	-0.001	0.004**	0.001	0.011***
	(0.006)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)
R^2	0.896	0.818	0.420	0.865	0.730	0.942
Observations	1581	1581	1581	1581	1581	1581

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by two years. Panel data OLS regressions with state and year fixed effects. Time span for innovation: 1976-2006. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 8: 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)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation	0.022***	0.026***	0.049***	0.060***	0.076***	0.094***
	(0.006)	(0.006)	(0.009)	(0.010)	(0.013)	(0.019)
Gdppc	0.032	0.050	0.063	0.055	0.060	0.046
	(0.028)	(0.036)	(0.044)	(0.055)	(0.068)	(0.095)
Popgrowth	0.553	0.618	1.089	1.595*	2.289**	3.307**
	(0.424)	(0.466)	(0.700)	(0.829)	(1.120)	(1.567)
Finance	0.066***	0.063**	0.109***	0.124***	0.079	0.021
	(0.020)	(0.025)	(0.036)	(0.046)	(0.072)	(0.106)
Government	-0.005	-0.009	-0.019*	-0.020*	-0.014	0.014
	(0.007)	(0.008)	(0.011)	(0.011)	(0.013)	(0.018)
Unemployment	-0.001	-0.005**	-0.006*	-0.008**	-0.010*	-0.014**
	(0.002)	(0.002)	(0.003)	(0.004)	(0.005)	(0.007)
TaxK	-0.018***	-0.023***	-0.039***	-0.047***	-0.061***	-0.084***
	(0.003)	(0.003)	(0.004)	(0.005)	(0.006)	(0.009)
TaxL	0.007*	0.012***	0.014**	0.018***	0.024***	0.031**
	(0.004)	(0.004)	(0.006)	(0.007)	(0.009)	(0.012)
R^2	0.818	0.877	0.896	0.893	0.891	0.864
Observations	1581	1581	1581	1581	1581	1581

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by two years. The dependent variables are taken in log. Panel data OLS regressions with state and year fixed effects. Time span for innovation: 1976-2006. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 9: Top 1% income share and innovation by entrants and incumbents

Dependent variable	Log of Top 1% Income Share					
	(1)	(2)	(3)	(4)	(5)	(6)
Measure of innovation	Patents	Patents	Patents	Cit5	Cit5	Cit5
Innovation by entrants	0.022***		0.018**	0.017***		0.014**
	(0.008)		(0.008)	(0.006)		(0.006)
Innovation by incumbents		0.018***	0.014**		0.025***	0.022***
		(0.007)	(0.007)		(0.006)	(0.006)
Gdppc	0.104**	0.082	0.089*	0.080	0.054	0.056
	(0.052)	(0.055)	(0.052)	(0.059)	(0.060)	(0.058)
Popgrowth	2.109***	1.997***	2.150***	2.287***	2.133***	2.208***
	(0.751)	(0.749)	(0.756)	(0.833)	(0.816)	(0.832)
Finance	0.096***	0.112***	0.104***	0.110***	0.135***	0.131***
	(0.032)	(0.032)	(0.032)	(0.033)	(0.033)	(0.033)
Government	-0.018	-0.024	-0.017	-0.023	-0.027	-0.020
	(0.021)	(0.021)	(0.021)	(0.020)	(0.021)	(0.021)
Unemployment	-0.002	-0.003	-0.003	-0.001	-0.003	-0.003
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
TaxK	-0.038***	-0.038***	-0.038***	-0.038***	-0.039***	-0.039***
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
TaxL	0.026***	0.026***	0.027***	0.022***	0.023***	0.024***
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
R^2	0.852	0.851	0.853	0.859	0.860	0.862
Observations	1530	1530	1530	1377	1377	1377

Notes: Variable description is given in Table 1. Innovation by entrants is a count of innovation that restricts to patents whose assignee first patented less than 3 years ago. Other patents enter in the count of Innovation by incumbents. Both these measures of innovation are taken in log and lagged by two years. Panel data OLS regressions with state and year fixed effects. Time span for innovation: 1980-2009 (columns 1 to 3) and 1980-2006 (columns 4 to 6). Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 10: Top 1% income share, innovation and the role of lobbying intensity

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Patents	Cit5	Claims
Innovation	0.905***	0.527**	0.837***	0.246	0.196	0.307*
	(0.000)	(0.014)	(0.000)	(0.172)	(0.312)	(0.091)
Innovation×Lobbying	-0.051***	-0.030**	-0.048***	-0.016	-0.011	-0.019*
	(0.000)	(0.015)	(0.000)	(0.132)	(0.320)	(0.073)
Lobbying	-0.305	-0.151	-0.095	0.053	-0.100	0.079
	(0.245)	(0.468)	(0.683)	(0.813)	(0.631)	(0.707)
Gdppc	0.107	0.014	0.105	0.095	-0.013	0.091
	(0.384)	(0.924)	(0.397)	(0.473)	(0.929)	(0.482)
Popgrowth	0.401	-0.146	0.379	0.640	0.150	0.622
	(0.738)	(0.897)	(0.754)	(0.613)	(0.902)	(0.622)
Finance	-0.021	-0.062	-0.027	-0.019	-0.057	-0.018
	(0.726)	(0.326)	(0.663)	(0.749)	(0.348)	(0.754)
Government	-0.107*	-0.189***	-0.108*	-0.117*	-0.221***	-0.115*
	(0.085)	(0.006)	(0.086)	(0.066)	(0.001)	(0.064)
Unemployment	-0.010**	-0.022***	-0.010**	-0.011**	-0.023***	-0.011**
	(0.026)	(0.000)	(0.023)	(0.016)	(0.000)	(0.015)
TaxK	-0.013**	-0.014**	-0.012**	-0.013**	-0.015***	-0.013**
	(0.025)	(0.012)	(0.028)	(0.023)	(0.010)	(0.022)
TaxL	-0.002	0.003	-0.003	-0.002	0.003	-0.002
	(0.840)	(0.815)	(0.815)	(0.844)	(0.811)	(0.884)
R^2	0.684	0.739	0.685	0.678	0.734	0.677
Observations	714	561	714	714	561	714

Notes: Lobbying is measured as explained in subsection 4.4. Other variable description is given in Table 1. Innovation is taken in log and lagged by two years. Columns 1 to 3 consider entrant innovation whether columns 4 to 6 consider incumbent innovations. The dependent variable is taken in log. Panel data OLS regressions with state and year fixed effects. Time span for innovation: 1996-2009 (columns 1, 3, 4 and 6) and 1996-2005 (columns 2 and 5). Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 11: Top 1% income share and innovation at different lags

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5
Lag of innovation	2 years	3 years	4 years	5 years	6 years	All lags
Innovation at $t-2$	0.043***					0.028
	(0.010)					(0.017)
Innovation at $t-3$		0.040***				0.017
		(0.009)				(0.015)
Innovation at $t-4$			0.039***			0.023
			(0.009)			(0.017)
Innovation at $t-5$				0.030***		-0.000
				(0.009)		(0.015)
Innovation at $t-6$					0.021**	-0.023
					(0.010)	(0.016)
Gdppc	0.035	0.035	0.033	0.044	0.057	0.029
	(0.062)	(0.062)	(0.062)	(0.062)	(0.061)	(0.063)
Popgrowth	2.212***	2.266***	2.305***	2.282***	2.283***	2.207***
	(0.840)	(0.839)	(0.827)	(0.829)	(0.828)	(0.847)
Finance	0.138***	0.134***	0.133***	0.126***	0.121***	0.140***
	(0.034)	(0.033)	(0.033)	(0.033)	(0.034)	(0.035)
Government	-0.025	-0.027	-0.028	-0.029	-0.030	-0.024
	(0.019)	(0.019)	(0.019)	(0.020)	(0.020)	(0.019)
Unemployment	-0.003	-0.003	-0.003	-0.002	-0.002	-0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
TaxK	-0.039***	-0.039***	-0.038***	-0.039***	-0.038***	-0.038***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
TaxL	0.022***	0.022***	0.022***	0.022***	0.021***	0.022***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
R^2	0.860	0.860	0.860	0.859	0.858	0.861
Observations	1377	1377	1377	1377	1377	1377

Notes: Variable description is given in Table 1. Innovation is taken in log. The lag between the dependent variable and the innovation measures ranges from 2 years to 6 years. Panel data OLS regressions with state and year fixed effects. Time span for innovation: 1980-2006. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 12: Descriptive statistics on the Senate appropriation committee composition.

	Number of years with			Number of years with	
	1 Senator	2 Senators		1 Senator	2 Senators
AK	28	0	**MT**	22	0
AL	14	0	**NC**	2	0
AR	29	0	**ND**	25	10
AZ	20	0	**NE**	17	0
CA	14	0	**NH**	32	0
CO	17	0	**NJ**	27	0
CT	12	0	**NM**	37	0
DE	3	0	**NV**	32	1
FL	21	0	**NY**	14	0
GA	10	0	**OH**	6	0
HI	33	6	**OK**	16	0
IA	20	2	**OR**	25	0
ID	24	0	**PA**	36	0
IL	12	0	**RI**	11	0
IN	9	0	**SC**	34	0
KS	7	0	**SD**	17	0
KY	26	0	**TN**	20	0
LA	33	0	**TX**	20	0
MA	8	0	**UT**	27	0
MD	29	1	**VA**	0	0
ME	3	0	**VT**	30	2
MI	1	0	**WA**	21	10
MN	0	0	**WI**	31	8
MO	30	0	**WV**	39	0
MS	31	8	**WY**	7	0

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 13: Senate Appropriation Committee composition and earmarks

Dependent variable	Log of earmarks				Cit5 univ	
	All earmarks		Research earmarks			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
SenateMember	0.401***	0.331***	0.330***	0.310***	0.096**	0.089**
	(0.076)	(0.074)	(0.113)	(0.103)	(0.048)	(0.037)
Gdppc	-1.003	-0.327	-4.092***	-3.182***	0.746**	0.448
	(0.708)	(0.703)	(1.002)	(0.988)	(0.367)	(0.283)
Ppgrowth	-0.805	-2.825	2.623	1.949	-4.337	-2.851
	(5.519)	(5.111)	(8.062)	(7.659)	(3.411)	(2.726)
Finance	0.651	0.213	0.292	0.290	-1.018***	-0.606***
	(0.479)	(0.422)	(0.601)	(0.524)	(0.253)	(0.203)
Government	-0.144	0.228	0.333	0.059	0.169	0.134
	(0.518)	(0.522)	(0.532)	(0.559)	(0.103)	(0.095)
Unemployment	-0.016	-0.016	-0.101*	-0.050	-0.052**	-0.030
	(0.037)	(0.031)	(0.055)	(0.054)	(0.022)	(0.018)
TaxK	0.050	0.085*	0.052	0.038	-0.025	-0.013
	(0.047)	(0.047)	(0.062)	(0.056)	(0.030)	(0.024)
TaxL	-0.062	-0.174*	-0.287**	-0.127	0.133***	0.083**
	(0.089)	(0.101)	(0.142)	(0.131)	(0.046)	(0.036)
$Y_{i,t-1}$		0.160***		0.201***		0.338***
		(0.039)		(0.040)		(0.031)
R^2	0.636	0.637	0.426	0.449	0.588	0.637
Observations	918	867	918	867	1428	1428

Notes: The dependent variable in columns 1 to 4 is the log of total earmarks received per capita in the state and comes from Cohen et al. (2011). Research earmarks have been selected based on the title on the appropriation bill. Columns 5 and 6 used the citations received within a 5 year window to patent assigned to universities. Panel data OLS regressions with state and year fixed effects. $Y_{i,t-1}$ denotes the lagged value of the dependent variable. Other variables description is given in Table 1. Autocorrelation and heteroskedasticity robust standard errors computed using the Newey-West variance estimator are reported in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 14: Regression of innovation on top 1% income share using instrument based on Appropriation Committee composition in the Senate

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.220**	0.185**	0.196**	0.218**	0.139**	0.144**
	(0.100)	(0.077)	(0.097)	(0.106)	(0.064)	(0.068)
Gdppc	-0.103	-0.079	-0.139	-0.122	-0.099	-0.071
	(0.107)	(0.091)	(0.133)	(0.123)	(0.105)	(0.098)
Popgrowth	1.960**	1.663*	2.223**	1.991**	1.553*	1.837*
	(0.923)	(0.953)	(1.014)	(0.928)	(0.937)	(0.950)
Finance	0.179***	0.213***	0.174***	0.192***	0.202***	0.210***
	(0.061)	(0.067)	(0.065)	(0.071)	(0.070)	(0.076)
Government	-0.097***	-0.078***	-0.089***	-0.091***	-0.038	-0.018
	(0.024)	(0.023)	(0.023)	(0.027)	(0.030)	(0.039)
Unemployment	-0.012**	-0.012**	-0.013**	-0.013**	-0.012**	-0.008*
	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.004)
TaxK	-0.040***	-0.039***	-0.040***	-0.042***	-0.040***	-0.035***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
TaxL	0.022***	0.016**	0.024***	0.025***	0.014**	0.013
	(0.008)	(0.007)	(0.009)	(0.010)	(0.007)	(0.008)
Highways	0.398	0.511	0.459	0.328	0.347	0.582
	(0.441)	(0.457)	(0.495)	(0.471)	(0.434)	(0.502)
Military	-0.002	-0.004	-0.002	-0.004	-0.008	-0.004
	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
R^2	0.906	0.910	0.854	0.851	0.844	0.823
F-stat on the excluded instruments	15.5	14.2	12.5	11.3	11.1	9.0
Observations	1700	1550	1700	1700	1550	1550

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by two years. 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 endogenous variable is set to 3 years. Time span for innovation: 1976-2009 for columns 1, 3 and 4 and 1976-2006 for columns 2, 5 and 6. DC is removed from the sample because it has no senators. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 15: 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	0.089***	0.016**			0.083***	0.018***
	(0.023)	(0.006)			(0.023)	(0.007)
Spillover			5.266***	0.735***	5.181***	0.754***
			(0.871)	(0.167)	(0.849)	(0.168)
Gdppc	1.039***	0.115**	0.911***	0.097	1.004***	0.088
	(0.160)	(0.046)	(0.181)	(0.061)	(0.171)	(0.067)
Popgrowth	-2.692	1.231*	-0.519	2.691***	0.601	2.688***
	(2.659)	(0.706)	(2.852)	(0.893)	(2.623)	(0.906)
Finance	-0.801***	0.072**	-0.666***	0.139***	-0.678***	0.133***
	(0.126)	(0.033)	(0.136)	(0.033)	(0.129)	(0.034)
Government	-0.037	-0.082***	-0.051	-0.102***	-0.031	-0.094***
	(0.064)	(0.024)	(0.082)	(0.029)	(0.077)	(0.028)
Unemployment	0.045***	-0.004	0.054***	0.000	0.059***	-0.001
	(0.012)	(0.003)	(0.011)	(0.004)	(0.011)	(0.004)
TaxK	0.016	-0.033***	0.013	-0.032***	0.006	-0.033***
	(0.014)	(0.004)	(0.017)	(0.005)	(0.017)	(0.005)
TaxL	-0.023**	0.005*	-0.023**	0.009**	-0.016	0.010***
	(0.009)	(0.003)	(0.011)	(0.004)	(0.011)	(0.003)
Highways	-2.546*	0.012			-5.553***	0.888**
	(1.433)	(0.317)			(1.325)	(0.382)
Military	0.003	-0.004			-0.006	-0.010
	(0.020)	(0.007)			(0.021)	(0.008)
R^2	0.926	0.927	0.852	0.867	0.858	0.869
F-stat	14.2	-	27.4	-	27.3	-
Observations	1550	1550	1350	1350	1350	1350

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 instrument and columns 5 and 6 use all instruments. The lags between the dependent variable and the instruments are set to match the corresponding second 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. Two additional controls for demand shocks are included, as explained in subsection 6.3, in columns 3 to 6. Time Span: 1976-2006 for columns 1 and 2 and 1981-2006 for columns 3 to 6. Variable description is given in Table 1. Panel data OLS regressions with state and year fixed effects. Innovation as well as the top 1% income share are taken in log. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 16: Robustness 1: regression of innovation on top 1% income share using two instruments

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of Innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.189***	0.156***	0.117***	0.116***	0.103***	0.139***
	(0.052)	(0.033)	(0.033)	(0.030)	(0.023)	(0.037)
Gdppc	-0.099	-0.074	-0.067	-0.045	-0.081	-0.080
	(0.089)	(0.076)	(0.083)	(0.076)	(0.081)	(0.082)
Popgrowth	2.923***	2.616***	2.971***	2.761***	2.598***	2.917***
	(1.001)	(1.004)	(0.947)	(0.904)	(0.981)	(1.011)
Finance	0.203***	0.243***	0.163***	0.167***	0.208***	0.255***
	(0.045)	(0.044)	(0.038)	(0.038)	(0.042)	(0.058)
Government	-0.110***	-0.089***	-0.111***	-0.116***	-0.051	-0.030
	(0.027)	(0.028)	(0.026)	(0.026)	(0.032)	(0.037)
Unemployment	-0.010***	-0.010***	-0.007**	-0.007**	-0.008**	-0.006
	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
TaxK	-0.034***	-0.033***	-0.032***	-0.032***	-0.034***	-0.030***
	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
TaxL	0.015***	0.012***	0.013***	0.013***	0.010***	0.011***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Highways	1.471***	1.757***	1.191***	1.165***	1.520***	1.847***
	(0.439)	(0.443)	(0.409)	(0.402)	(0.435)	(0.512)
Military	-0.006	-0.010	-0.007	-0.009	-0.013	-0.005
	(0.007)	(0.008)	(0.007)	(0.007)	(0.008)	(0.009)
R^2	0.836	0.851	0.844	0.850	0.831	0.765
F-stat on excluded instruments	19.3	27.3	33.7	47.0	28.6	14.8
Sargan-Hansen J-stat (p-value)	0.290	0.447	0.168	0.171	0.398	0.979
Observations	1500	1350	1500	1500	1350	1350

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by two years. 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 section6.3. The lag between the first instrument and the endogenous variable is set to 3 years while the lag between the second instrument and the endogenous variable is 1 year. Two additional controls for demand shocks are included, as explained in subsection 6.3. Time span: 1983-2011 for columns 1 1983-2008 for columns 2 to 6. DC is removed from the sample because it has no senators. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 17: Robustness 2: financial sector and natural resources

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation	0.058***	0.057***	0.049***	0.031***	0.053***	0.049***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Gdppc	-0.107**	0.026	0.067	0.058	0.148***	0.068
	(0.014)	(0.545)	(0.130)	(0.163)	(0.000)	(0.123)
Popgrowth	1.072	1.107	1.117	1.064	1.266*	1.116
	(0.139)	(0.121)	(0.109)	(0.121)	(0.053)	(0.109)
Finance	0.050	0.153***	0.114***	0.110***	0.143***	0.113***
	(0.132)	(0.001)	(0.001)	(0.001)	(0.000)	(0.002)
Government	0.017	-0.015	-0.021*	-0.014	-0.015	-0.020*
	(0.140)	(0.151)	(0.060)	(0.188)	(0.147)	(0.077)
Unemployment	-0.012***	-0.010***	-0.007**	-0.008***	-0.006*	-0.007**
	(0.000)	(0.003)	(0.041)	(0.009)	(0.088)	(0.043)
TaxK	-0.035***	-0.031***	-0.036***	-0.037***	-0.029***	-0.036***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TaxL	0.006**	0.003	0.007**	0.008**	0.010***	0.007**
	(0.049)	(0.344)	(0.023)	(0.014)	(0.001)	(0.024)
RemunFinance	0.339***					
	(0.000)					
EFD				0.661***		
				(0.000)		
Mining+Oil					0.036***	
					(0.000)	
R^2	0.905	0.896	0.897	0.900	0.903	0.896
Observations	1581	1457	1581	1581	1581	1581

Notes: Variable *Mining+oil* measure the share of oil related and natural resources extraction activities in GDP, variable *RemunFinance* measures the compensation per employee in the financial sector and variable *EFD* measures the financial dependence of innovation. Other variables description is given in Table 1. Innovation is taken in log and lagged by two years. Column 1 controls for average compensation in the financial sector, column 2 drops NY, CT, DE and MA (the state with the largest financial sectors), column 3 removes finance-related patents, column 4 controls for financial dependence in the state as explained in section 6.1, column 5 controls for the size of oil and mining sectors and column 6 removes oil-related patents from the count of citations. Time Span: 1976-2008. Panel data OLS regressions with state and year fixed effects. Innovation as well as the top 1% income share are taken in log. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 18: Simulation results

Moments			Parameters
Definition and source	Target	Simulations	
M_1Average top 1% share (own data)	0.130	0.13	η_L= 1.16
M_2 Ratio of entrant to incumbent citations (own data)	0.2	0.25	η_H= 1.35
M_3 Elasticity of top 1% w.r.t innovation (Garcia-Macia et al., 2016)	0.185	0.184	θ_I= 0.7
M_4 Average mark-up (Jaimovich and Floetotto, 2008)	1.2	1.20	θ_E= 7.3
M_5 Entrant share of employment (own data)	0.03	0.031	$L = 74.8$
M_6 Growth rate	0.02	0.020	$\phi = 0.196$

Notes: Definition and value of the targeted moments, average value for the targeted moments in 500 draws of simulated data and parameters

Table 19: Top 1% income share and innovation - CZ level panel

Dependent variable	Log of Top 1% Income Share				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Measure of innovation	Patents	Patents	Patents	Patents	Patents
Innovation	0.021*	0.019*	0.019*	0.019*	0.018*
	(0.012)	(0.011)	(0.011)	(0.011)	(0.010)
Gdppc		-0.352	-0.359	-0.540**	-0.596**
		(0.217)	(0.217)	(0.260)	(0.288)
Popgrowth			0.333	0.277	0.011
			(0.561)	(0.508)	(0.428)
Finance				0.002	0.007
				(0.086)	(0.088)
Government				-0.187**	-0.166**
				(0.088)	(0.078)
Unemployment					-1.814
					(1.452)
R^2	0.816	0.818	0.812	0.813	0.814
Observations	5599	5599	5571	5570	5570

Notes: Variable description is given in Table 1. A dummy equal to one if the CZ belongs to an urban area is included but not reported. Panel fixed effect regression with CZs weighted by population and state×year dummies. Time span for innovation: 1998 and 2003-2009. Regressions also include a dummy for being an urban CZ. Heteroskedasticity robust standard errors clustered at the state level are reported in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 20: 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)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation	0.015	0.076*	0.028	0.023**	0.112**	0.053**	0.012
	(0.010)	(0.042)	(0.025)	(0.010)	(0.042)	(0.025)	(0.017)
Gdppc	0.025	0.416*	0.158	-0.074	0.007	-0.144	-0.051
	(0.054)	(0.235)	(0.136)	(0.062)	(0.255)	(0.148)	(0.106)
Popgrowth	-1.156	-1.322	-5.852**	-1.944**	-4.976	-8.218***	-7.210***
	(0.850)	(3.667)	(2.539)	(0.838)	(3.628)	(2.288)	(1.600)
Government	0.047	0.263*	0.119	0.038	0.227	0.088	0.051
	(0.032)	(0.133)	(0.090)	(0.033)	(0.138)	(0.093)	(0.066)
Finance	0.032	0.035	0.093*	0.016	-0.023	0.045	0.046
	(0.021)	(0.083)	(0.054)	(0.019)	(0.073)	(0.054)	(0.039)
Unemployment	-0.025	0.720	-0.202	-0.201	-0.026	-0.740	-0.723*
	(0.212)	(0.908)	(0.604)	(0.211)	(0.872)	(0.550)	(0.398)
Tax	0.000	0.001	0.001	-0.001	-0.004	-0.003	-0.001
	(0.001)	(0.006)	(0.004)	(0.002)	(0.006)	(0.005)	(0.004)
School Expenditure				0.008	0.027	0.024	0.016
				(0.009)	(0.034)	(0.024)	(0.019)
Employment Manuf				-0.391***	-1.682***	-1.177***	-0.720***
				(0.110)	(0.401)	(0.332)	(0.247)
R^2	0.146	0.180	0.168	0.197	0.225	0.218	0.264
Observations	666	674	674	662	670	670	670

Notes: Variable description is given in Table 1. The number of citations per inhabitants is averaged over the period 1998-2008 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. Cross section OLS regressions with CZs weighted by population. Regressions also include a dummy for being an urban CZ. Heteroskedasticity robust standard errors clustered at the state level are reported in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 21: 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)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation by entrants	0.023**	0.111***	0.048**				0.019*
	(0.009)	(0.039)	(0.022)				(0.010)
Innovation by incumbents				0.016**	0.075**	0.034*	0.006
				(0.008)	(0.033)	(0.020)	(0.007)
Gdppc	-0.081	-0.021	-0.137	-0.048	0.145	-0.072	-0.086
	(0.057)	(0.235)	(0.143)	(0.064)	(0.270)	(0.146)	(0.058)
Popgrowth	-1.770**	-4.074	-7.770***	-1.849**	-4.476	-7.948***	-1.825**
	(0.821)	(3.550)	(2.222)	(0.838)	(3.670)	(2.301)	(0.863)
Finance	0.018	-0.015	0.049	0.017	-0.021	0.046	0.018
	(0.018)	(0.070)	(0.053)	(0.019)	(0.073)	(0.054)	(0.019)
Government	0.035	0.210	0.081	0.039	0.231	0.090	0.035
	(0.033)	(0.136)	(0.094)	(0.034)	(0.145)	(0.096)	(0.033)
Unemployment	-0.225	-0.141	-0.805	-0.199	-0.028	-0.747	-0.203
	(0.208)	(0.866)	(0.549)	(0.217)	(0.900)	(0.564)	(0.210)
Tax	-0.001	-0.003	-0.003	-0.001	-0.004	-0.003	-0.001
	(0.002)	(0.006)	(0.005)	(0.002)	(0.006)	(0.005)	(0.002)
School Expenditure	0.009	0.035	0.028	0.007	0.024	0.023	0.009
	(0.009)	(0.033)	(0.024)	(0.009)	(0.036)	(0.025)	(0.009)
Employment Manuf	-0.334***	-1.400***	-1.037***	-0.385***	-1.640***	-1.146***	-0.358***
	(0.109)	(0.394)	(0.323)	(0.113)	(0.413)	(0.339)	(0.113)
R^2	0.197	0.225	0.214	0.185	0.207	0.209	0.201
Observations	662	670	670	662	670	670	662

Notes: Variable description is given in Table 1. The number of citations per inhabitants is averaged over the period 1998-2008 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. Cross section OLS regressions with CZs weighted by population. Regressions also include a dummy for being an urban CZs. Heteroskedasticity robust standard errors clustered at the state level are reported in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 22: Innovation and social mobility at the CZ level. Entrants and incumbents innovation and the role of lobbying

Dependent variable	AM25			
	(1)	**(2)**	**(3)**	**(4)**
Innovation by entrants	0.001	0.038***	0.001	0.035***
	(0.008)	(0.009)	(0.007)	(0.012)
Innovation by incumbents			0.001	0.005
			(0.006)	(0.008)
Gdppc	-0.055	-0.085	-0.056	-0.087
	(0.108)	(0.053)	(0.107)	(0.053)
Popgrowth	-3.423**	-0.836	-3.431**	-0.885
	(1.289)	(0.924)	(1.301)	(0.968)
Finance	0.031	0.015	0.031	0.014
	(0.025)	(0.020)	(0.025)	(0.021)
Government	-0.020	0.037	-0.020	0.037
	(0.022)	(0.040)	(0.023)	(0.040)
Unemployment	-0.899**	7-0.045	-0.891**	0.056
	(0.333)	(0.217)	(0.337)	(0.217)
Tax	0.001	-0.001	0.001	-0.001
	(0.002)	(0.001)	(0.002)	(0.001)
School Expenditure	0.015	0.013	0.015	0.013
	(0.009)	(0.009)	(0.010)	(0.009)
Employment Manuf	-0.298***	-0.283**	-0.302***	-0.303**
	(0.105)	(0.123)	(0.111)	(0.126)
R^2	0.408	0.268	0.408	0.271
Observations	331	331	331	331

Notes: Variable description is given in Table 1. The number of citations per inhabitants is averaged over the period 1998-2008 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. Columns 1 and 3 restrict to CZs above median in terms of lobbying intensity, where lobbying is measured as explained in subsection 8.2 while columns 2 and 4 restrict to CZs below median. Cross section OLS regressions with CZs weighted by population. Regressions also include a dummy for being an urban CZ. Heteroskedasticity robust standard errors clustered at the state level are reported in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Online Appendix for
"Innovation and Top Income Inequality "
Not for publication unless requested

A Theoretical appendix

A.1 Proofs 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.

A.2 Entrepreneurial share of income net of innovation costs

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

$$\begin{aligned}
& rel_net_share \qquad (16)\\
= & \left(Entrepreneur_share_t - \theta_E \frac{x_E^2}{2} - \theta_I \frac{x_I^2}{2}\right) / \left(\frac{w_t M}{Y_t(1+L)} L\right) \\
= & \left(\pi_L + \frac{\pi_H - \pi_L}{2} x_I^* + \left(\frac{\pi_H}{2} + \frac{w_t M}{2Y_t(1+L)} - \pi_L\right)(1-z)\, x_E^*\right) / \left(\frac{w_t M L}{Y_t(1+L)}\right)
\end{aligned}$$

where we used (5), (7), the equilibrium values (9) and (10) and took into account the fact that there are $M/(1+L)$ product lines. This expression shows that a higher rate of incumbent innovation will raise the net entrepreneur share of income, whereas a higher rate of entrant innovation will only raise the net entrepreneurial share of income if $\frac{1}{2}\frac{1+L}{M}\pi_H+\frac{1}{2}\frac{w_t}{Y_t}-\frac{1+L}{M}\pi_L>0$ (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 $\frac{1+L}{M}\pi_H Y_t-w_t$ by innovating but she destroys the rents $\frac{1+L}{M}\pi_L Y_t$ of the incumbent. Formally, we can show:

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).

Proof of Proposition 3 Using (6), we rewrite:

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

We then obtain

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

Using (16), we get:

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

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

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 $\frac{1}{2}\pi_H + \frac{1}{2}\frac{w_t M}{Y_t(1+L)} - \pi_L > 0$.

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 $\frac{1}{2}\pi_H + \frac{1}{2}\frac{w_t M}{Y_t(1+L)} - \pi_L > 0$. An increase in θ_I 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

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

Note that $x_E^* < 1$, requires $\left(\pi_H - \frac{w_t M}{Y_t(1+L)}\right)(1-z) < \theta_E$. Moreover as $L > 1$, we must have

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

Hence the relative net share is always decreasing in θ_I.

Finally consider the case where L is large such that $\frac{w_t M}{Y_t(1+L)}$ is small then we have

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

therefore

$$
\begin{aligned}
&\frac{\partial rel\_net\_share}{\partial x_E^*} \\
\approx\ & \left(\left(\frac{1}{2}\pi_H - \pi_L\right)\frac{w_t M}{Y_t(1+L)} + \frac{(\pi_H - \pi_L)}{L}\begin{pmatrix} \pi_L + \frac{1}{2}(\pi_H - \pi_L)\, x_I^* \\ + \left(\frac{1}{2}\pi_H - \pi_L\right)(1-z)\, x_E^* \end{pmatrix}\right)\left(\frac{Y_t(1+L)}{w_t M}\right)^2 \frac{1-z}{L} \\
\approx\ & \begin{pmatrix} \left(\frac{1}{2}\pi_H - \pi_L\right)\left(1 - \pi_L - (\pi_H - \pi_L)\left(\frac{\pi_H - \pi_L}{\theta_I} + (1-z)\frac{\pi_H}{\theta_E}\right)\right) \\ + (\pi_H - \pi_L)\left(\pi_L + \frac{1}{2}\frac{(\pi_H - \pi_L)^2}{\theta_I} + \left(\frac{1}{2}\pi_H - \pi_L\right)(1-z)\frac{\pi_H}{\theta_E}\right) \end{pmatrix}\left(\frac{Y_t(1+L)}{w_t M L}\right)^2 (1-z) \\
\approx\ & \left(\left(\frac{1}{2}\pi_H - \pi_L\right)(1-\pi_L) + (\pi_H - \pi_L)\,\pi_L + \frac{1}{2}\frac{\pi_L(\pi_H - \pi_L)^2}{\theta_I}\right)\left(\frac{Y_t(1+L)}{w_t M L}\right)^2 (1-z)
\end{aligned}
$$

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.3 Proofs for subsection 2.3.2

From (9), we have: $\frac{\partial x_I^*}{\partial \eta_L} = -\frac{1}{\eta_L^2}\frac{1}{\theta_I} < 0$, whereas:

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

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:

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

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.4 Shared rents

A.4.1 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. For simplicity we assume that $M = 1 + L$, so that the economy is 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:

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

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:

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

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.[58]

[58]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:

$$Top\_share = \pi_H \mu^* + \pi_L (1-\mu^*) - \frac{\theta_E x_E^2}{2} - \frac{\theta_I x_I^2}{2}, \tag{17}$$

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 (17) 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.4.2 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. For simplicity we assume that $M=1+L$. 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, and 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.[59] 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

[59] 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

$$\frac{w}{Y} = v\widehat{h}. \tag{18}$$

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

$$x_I^* = \frac{\pi_H - \pi_L}{\theta v} \text{ and } x_E^* = \pi_H \frac{1-z}{\theta v}, \tag{19}$$

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:

$$\begin{aligned} \theta\left(\frac{x_I^{*2}}{2} + \frac{x_E^{*2}}{2}\right) &= L\int_{\widehat{h}}^{\overline{h}} h dh \Leftrightarrow \\ (\pi_H - \pi_L)^2 + \pi_H^2 (1-z)^2 &= \theta v^2 L \frac{\overline{h}^2 - \widehat{h}^2}{\overline{h}}. \end{aligned} \tag{20}$$

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 (18) 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:

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

Plugging (21) into (20), we obtain that the wage rate for innovative human capital is uniquely defined by:

$$vL\overline{h} = 1 - \pi_L + \pi_L \pi_H \frac{(1-z)^2}{\theta v}. \tag{22}$$

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:

$$Innov\_share = \int_{\widehat{h}}^{\overline{h}} vLhdh = vL\left(\overline{h}^2 - \widehat{h}^2\right) / \left(2\overline{h}\right). \tag{23}$$

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:

$$\begin{aligned} Owner\_share &= \pi_L(1-\mu) + \pi_H\mu - Innov\_share \\ &= \pi_L + \frac{1}{2\theta v}\left((\pi_H - \pi_L)^2 + (\pi_H - 2\pi_L)\pi_H(1-z)^2\right). \end{aligned} \tag{24}$$

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 (22) 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 (19) we then obtain that both entrant innovation x^* and incumbent innovation x_I^* decrease with θ. Differentiating (20) 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 (22) to obtain the latter equality.

Using (20) in (23), 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 (19) in (24) 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.

A.5 CES production technology

For simplicity we assume that $M = 1 + L$, and we change the production function to:

$$Y_t = \left(\int_0^1 y_{it}^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}, \tag{25}$$

with $y_{it} = q_{it} l_{it}$ and $\sigma > 1$. A competitive fringe has access (at the beginning of the period) to productivity level q_{it}/η_L. We focus only on productive innovations here, so that an when a firm innovates, q_{it} increases by a factor η_H and the fringe gets access to (the previous) q_{it}. We assume that η_H is small enough that the firm is forced to limit pricing.

We take the final good as the numeraire. Then we get that the equilibrium prices are:

$$p_{it}^L = \eta_L \frac{w_t}{q_{it}} \text{ in sectors without innovation} \tag{26}$$

$$p_{it}^H = \eta_H \frac{w_t}{q_{it}} \text{ in sectors with innovation} \tag{27}$$

Moreover $y_{it} = p_{it}^{-\sigma} Y_t$ so that $\pi_{it} = \left(p_{it} - \frac{w_t}{q_{it}} \right) p_{it}^{-\sigma} Y_t$. Hence:

$$\pi_{it}^L = \frac{\eta_L - 1}{\eta_L^\sigma} \left(\frac{w_t}{q_{it}} \right)^{1-\sigma} Y_t = \frac{\eta_L - 1}{\eta_L^\sigma} \left(\frac{w_t}{q_{it}^0} \right)^{1-\sigma} Y_t,$$

$$\pi_{it}^H = \frac{\eta_H - 1}{\eta_H^\sigma} \left(\frac{w_t}{q_{it}} \right)^{1-\sigma} Y_t = \frac{\eta_H - 1}{\eta_H} \left(\frac{w_t}{q_{it}^0} \right)^{1-\sigma} Y_t.$$

Here the superscript "0" indicates productivities pre-innovation.

A natural assumption to make (e.g. see Aghion and Howitt, 1998, Chapter 9) is that pre-innovation, all agents in the economy have access to the technology $q_{it}^0 = \int_0^1 q_{i(t-1)} di = Q_{t-1}$. Then

$$\Pi_t^L = \frac{\eta_L - 1}{\eta_L^\sigma} \left(\frac{w_t}{Q_{t-1}} \right)^{1-\sigma} Y_t \text{ and } \Pi_t^H = \frac{\eta_H - 1}{\eta_H} \left(\frac{w_t}{Q_{t-1}} \right)^{1-\sigma} Y_t$$

so that $\Pi_t^H > \Pi_t^L$.

Note that $\Pi_t^H / \Pi_t^L = \frac{\eta_H - 1}{\eta_H} / \frac{\eta_L - 1}{\eta_L^\sigma}$ is bigger than in the Cobb-Douglas case. Another difference with the Cobb-Douglas case is the term $\left(\frac{w_t}{Q_{t-1}} \right)^{1-\sigma}$ which reflects a competition effect whereby innovation by others increases the wage and therefore raises the price of my own good because the production cost has increased.

To express the entrepreneur share of income, we need to solve for the equilibrium wage

w_t. (26), (27) and the dynamics of q_{it} give

$$p_t^L = \eta_L \frac{w_t}{q_{t-1}} \text{ and } p_t^H = \frac{w_t}{q_{t-1}},$$

which together with the price normalization

$$\mu_t p_{t,H}^{1-\sigma} + (1-\mu_t)\, p_{t,L}^{1-\sigma} = 1,$$

immediately yields:

$$\mu_t \left(\frac{w_t}{q_{t-1}}\right)^{1-\sigma} + (1-\mu_t)\left(\eta_L \frac{w_t}{q_{t-1}}\right)^{1-\sigma} = 1,$$

so that:

$$\left(\frac{w_t}{q_{t-1}}\right)^{1-\sigma} = \frac{1}{\mu_t + (1-\mu_t)\,\eta_L^{1-\sigma}}. \tag{28}$$

The entrepreneur share of income can then be written as:

$$\begin{aligned} entrepreneur\_share_t &= \frac{\mu_t \Pi_{H,t} + (1-\mu_t)\,\Pi_{L,t}}{Y_t} \qquad (29)\\ &= 1 - \frac{1}{\eta_L} + \left(\frac{1}{\eta_L} - \frac{1}{\eta_H}\right)\frac{\mu_t}{\mu_t + (1-\mu_t)\,\eta_L^{1-\sigma}} \end{aligned}$$

This expression is increasing in μ_t, which is still given by $\mu_t = x_{It} + (1-z)\,x_{Et}$. In addition, we know that social mobility is still equal to $\Psi_t = x_{Et}\,(1-z)\,/L$. Therefore, we still get:

Proposition 6 *(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.*

We now turn to the endogenous determination of the innovation rates of entrants and incumbents. We use the same innovation function as in the baseline model. The maximization problem of the incumbent is:

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

We then obtain that the optimal innovation decision is simply

$$x_{I,t} = x_I^* = \frac{1}{\theta_I}\left(1 - \frac{1}{\eta_H} - \eta_L^{1-\sigma}\left(1 - \frac{1}{\eta_L}\right)\right)\left(\frac{w_t}{q_{t-1}}\right)^{1-\sigma}. \tag{30}$$

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

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

Therefore, we get

$$x_{E,t} = x_E^* = \left(\frac{\eta_H - 1}{\eta_H}\left(\frac{w_t}{q_{t-1}}\right)^{1-\sigma} - \frac{w_t}{Y_t}\right)\frac{(1-z)}{\theta_E}. \tag{31}$$

Using (29), we get:

$$\frac{w_t}{Y_t} = \frac{1 - entrepreneur\_share_t}{L} = \frac{1}{L}\frac{\mu_t\frac{1}{\eta_H} + (1-\mu_t)\,\eta_L^{1-\sigma}\frac{1}{\eta_L}}{\mu_t + (1-\mu_t)\,\eta_L^{1-\sigma}}.$$

Plugging this expression and (28) into (30) and (31) we obtain:

$$x_{I,t} = \frac{1}{\theta_I}\frac{1 - \frac{1}{\eta_H} - \eta_L^{1-\sigma}\left(1 - \frac{1}{\eta_L}\right)}{\mu_t\left(1 - \eta_L^{1-\sigma}\right) + \eta_L^{1-\sigma}}, \tag{32}$$

$$x_{E,t} = x_E^* = \frac{1 - \frac{1}{\eta_H} - \frac{1}{L}\left[\mu_t\frac{1}{\eta_H} + (1-\mu_t)\,\eta_L^{1-\sigma}\frac{1}{\eta_L}\right]}{\mu_t + (1-\mu_t)\,\eta_L^{1-\sigma}}\frac{1-z}{\theta_E}. \tag{33}$$

The above expression shows that a change in μ_t (for instance because of a change in incumbent innovation x_I) has an ambiguous effect on entrant innovation x_{Et}. On the one hand as in the Cobb-Douglas case, an increase in μ_t reduces w_t/Y_t and therefore makes the outside option of the entrant less appealing, which leads to higher innovation by the entrant. On the other hand, an increase in μ_t also increases w_t/q_{t-1}. This is the competition effect mentioned above which decreases entrant innovation. As a result, a reduction in incumbent innovation costs (θ_I), which increases incumbent innovation may reduces entrant innovation and thereby social mobility. Overall, we obtain:

Proposition 7 *An increase in entrant innovation costs θ_E reduces entrant innovation x_E, incumbent innovation x_I and social mobility. An increase in incumbent innovation costs θ_I reduces incumbent innovation x_I and total innovation μ but has an ambiguous impact on* *entrant innovation and social mobility.*

Proof. To solve for the the total number of innovations, combine (32) and (33) to get:

$$\mu_t = \frac{1}{\theta_I}\frac{1-\frac{1}{\eta_H}-\eta_L^{1-\sigma}\left(1-\frac{1}{\eta_L}\right)}{\mu_t\left(1-\eta_L^{1-\sigma}\right)+\eta_L^{1-\sigma}} + \frac{1-\frac{1}{\eta_H}-\frac{1}{L}\left[\mu_t\frac{1}{\eta_H}+\left(1-\mu_t\right)\eta_L^{1-\sigma}\frac{1}{\eta_L}\right]}{\mu_t+\left(1-\mu_t\right)\eta_L^{1-\sigma}}\frac{\left(1-z\right)^2}{\theta_E}.$$

This expression can be rewritten as:

$$\frac{1}{\theta_I}\frac{1-\frac{1}{\eta_H}-\eta_L^{1-\sigma}\left(1-\frac{1}{\eta_L}\right)}{\left(\mu_t\left(1-\eta_L^{1-\sigma}\right)+\eta_L^{1-\sigma}\right)\mu_t} + \frac{\left(1-\frac{1}{\eta_H}-\frac{1}{L}\frac{1}{\eta_L^{\sigma}}\right)\frac{1}{\mu_t}-\frac{1}{L}\left[\frac{1}{\eta_H}-\frac{1}{\eta_L^{\sigma}}\right]}{\mu_t+\left(1-\mu_t\right)\eta_L^{1-\sigma}}\frac{\left(1-z\right)^2}{\theta_E}=1 \tag{34}$$

As $1-\frac{1}{\eta_H}-\frac{1}{L}\frac{1}{\eta_L^{\sigma}}>0$ (which results from assuming that for any μ_t, $\Pi_H > w$), the LHS is decreasing in μ_t. Therefore this expression defines μ_t uniquely, and we get that total innovation μ_t decreases in the entrant and incumbent innovation costs θ_E and θ_I.

Since μ is decreasing in θ_E and x_I is decreasing in μ, we have that x_I is increasing in θ_E. Assume by contradiction that x_E is also increasing in θ_E, then μ is increasing in θ_E which is impossible. Therefore x_{Et} and so Ψ (social mobility) are decreasing in θ_E.

Rewrite (34) as:

$$\frac{x_{It}}{\mu_t} + \frac{\left(1-\frac{1}{\eta_H}-\frac{1}{L}\frac{1}{\eta_L^{\sigma}}\right)\frac{1}{\mu_t}-\frac{1}{L}\left[\frac{1}{\eta_H}-\frac{1}{\eta_L^{\sigma}}\right]}{\mu_t+\left(1-\mu_t\right)\eta_L^{1-\sigma}}\frac{\left(1-z\right)^2}{\theta_E}=1, \tag{35}$$

an increase in θ_I decreases μ_t which increases the LHS therefore it must decrease x_{It}. Hence x_{It} is decreasing in θ_I. Yet, since x_{Et} is ambiguous in μ, it is also ambiguous in θ_I, and so is Ψ_t. ■

B Additional empirical results

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

Dependent variable	Log of Top 0.1% Income Share			Log of Top 0.01% Income Share		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation by entrants	0.030***		0.025**	0.037***		0.030**
	(0.010)		(0.010)	(0.014)		(0.014)
Innovation by incumbents		0.041***	0.036***		0.059***	0.053***
		(0.009)	(0.009)		(0.012)	(0.012)
Gdppc	0.063	0.024	0.025	0.048	-0.010	-0.008
	(0.094)	(0.097)	(0.095)	(0.131)	(0.133)	(0.131)
Popgrowth	4.047***	3.823***	3.920***	5.491***	5.169***	5.305***
	(1.382)	(1.361)	(1.378)	(1.923)	(1.895)	(1.916)
Finance	0.139***	0.178***	0.173***	0.139**	0.195***	0.189***
	(0.048)	(0.050)	(0.050)	(0.064)	(0.068)	(0.067)
Government	-0.017	-0.025	-0.014	0.011	0.002	0.016
	(0.032)	(0.032)	(0.032)	(0.041)	(0.041)	(0.041)
Unemployment	-0.007	-0.010*	-0.009*	-0.011	-0.016**	-0.015*
	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)
TaxK	-0.057***	-0.058***	-0.059***	-0.074***	-0.076***	-0.076***
	(0.008)	(0.007)	(0.007)	(0.010)	(0.010)	(0.010)
TaxL	0.026***	0.027***	0.029***	0.028**	0.029**	0.032**
	(0.010)	(0.010)	(0.010)	(0.013)	(0.013)	(0.013)
R^2	0.826	0.827	0.829	0.785	0.787	0.789
Observations	1377	1377	1377	1377	1377	1377

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by two years. Panel data OLS regressions with state and year fixed effects. Time span for innovation: 1980-2006. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B2: Distribution of patents located by inventors and assignees

State	Inventors	Assignees	State	Inventors	Assignees
AK	0.1%	0.0%	**MT**	0.1%	0.1%
AL	0.4%	0.3%	**NC**	2.3%	1.4%
AR	0.2%	0.1%	**ND**	0.1%	0.0%
AZ	1.8%	0.7%	**NE**	0.2%	0.2%
CA	22.8%	24.0%	**NH**	0.7%	0.4%
CO	2.2%	1.1%	**NJ**	4.1%	5.1%
CT	2.0%	2.9%	**NM**	0.4%	0.3%
DC	0.1%	1.2%	**NV**	0.4%	0.6%
DE	0.4%	1.9%	**NY**	7.1%	11.6%
FL	2.8%	1.9%	**OH**	3.5%	3.8%
GA	1.7%	1.0%	**OK**	0.5%	0.3%
HI	0.1%	0.0%	**OR**	1.8%	0.7%
IA	0.7%	0.7%	**PA**	3.5%	2.9%
ID	1.8%	2.2%	**RI**	0.3%	0.2%
IL	3.9%	5.0%	**SC**	0.6%	0.4%
IN	1.5%	0.9%	**SD**	0.1%	0.0%
KS	0.5%	0.3%	**TN**	0.8%	0.6%
KY	0.5%	0.4%	**TX**	7.0%	6.6%
LA	0.5%	0.2%	**UT**	0.8%	0.6%
MA	4.4%	4.1%	**VA**	1.4%	0.9%
MD	1.7%	1.1%	**VT**	0.4%	0.1%
ME	0.2%	0.1%	**WA**	2.9%	2.6%
MI	4.2%	4.8%	**WI**	2.0%	1.8%
MN	3.0%	3.0%	**WV**	0.1%	0.0%
MO	1.0%	0.8%	**WY**	0.1%	0.0%
MS	0.2%	0.1%			

Notes: Distribution granted patent with an application year equal to 2000 by state whether the allocation is based on the address of the inventors or the address of the assignees.

Table B3: Top 1% income share and innovation allocation by assignee

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.048***	0.031***	0.024***	0.023***	0.021***	0.023***
	(0.010)	(0.006)	(0.006)	(0.006)	(0.005)	(0.004)
Gdppc	0.062	0.081*	0.094**	0.093**	0.072	0.077*
	(0.042)	(0.047)	(0.042)	(0.045)	(0.044)	(0.044)
Popgrowth	1.138*	0.977	0.964	0.963	0.962	0.985
	(0.673)	(0.697)	(0.654)	(0.672)	(0.690)	(0.677)
Finance	0.088**	0.089***	0.074**	0.075**	0.086**	0.087**
	(0.034)	(0.034)	(0.034)	(0.034)	(0.035)	(0.035)
Government	-0.019*	-0.027**	-0.022**	-0.023**	-0.024**	-0.025**
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Unemployment	-0.007**	-0.006*	-0.006*	-0.006*	-0.006*	-0.006*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
TaxK	-0.024***	-0.024***	-0.024***	-0.024***	-0.024***	-0.025***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
TaxL	0.013***	0.012***	0.013***	0.013***	0.011***	0.011***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
R^2	0.892	0.897	0.891	0.891	0.896	0.896
Observations	1734	1581	1734	1734	1581	1581

Notes: Innovation is taken in log and lagged by two years and is assigned to a state using the assignee location. Panel data OLS regressions with state and year fixed effects. Time span for innovation: 1976-2009 (columns 1, 3 and 4) and 1976-2006 (columns 2, 5 and 6). Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B4: Top 1% income share and innovation with clustered standard errors

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.032	0.049***	0.018	0.023	0.024***	0.019***
	(0.020)	(0.015)	(0.014)	(0.016)	(0.007)	(0.006)
Gdppc	0.078	0.051	0.086	0.085	0.064	0.075
	(0.086)	(0.085)	(0.087)	(0.087)	(0.084)	(0.089)
Popgrowth	0.893	1.042	0.885	0.867	0.937	1.022
	(0.627)	(0.650)	(0.619)	(0.623)	(0.631)	(0.629)
Finance	0.082	0.113*	0.074	0.079	0.097*	0.093
	(0.061)	(0.060)	(0.062)	(0.061)	(0.058)	(0.059)
Government	-0.018	-0.020	-0.018	-0.019	-0.018	-0.017
	(0.027)	(0.025)	(0.027)	(0.027)	(0.025)	(0.025)
Unemployment	-0.006	-0.007	-0.006	-0.006	-0.006	-0.005
	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Tax K	-0.024***	-0.025***	-0.024***	-0.024***	-0.025***	-0.024***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Tax L	0.013**	0.011*	0.012*	0.012*	0.011*	0.010*
	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.006)
R^2	0.890	0.897	0.890	0.890	0.896	0.895
Observations	1734	1581	1734	1734	1581	1581

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by two years. The dependent variable is the log of the top 1% income share. Panel data OLS regressions with state and year fixed effects. Time span for innovation: 1976-2009 (columns 1, 3 and 4) and 1976-2006 (columns 2, 5 and 6). Heteroskedasticity robust standard errors clustered at the state level are presented in parenthesis. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B5: Top 1% income share and innovation by entrants and incumbents - alternative definition of entrants

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Patents	Patents	Cit5	Cit5	Cit5
Innovation by entrants	0.013		0.006	0.020***		0.016**
	(0.009)		(0.009)	(0.007)		(0.007)
Innovation by incumbents		0.021***	0.020***		0.026***	0.023***
		(0.007)	(0.007)		(0.005)	(0.006)
Gdppc	0.124**	0.076	0.100*	0.085	0.048	0.055
	(0.053)	(0.055)	(0.053)	(0.059)	(0.060)	(0.059)
popgrowth	2.094***	2.025***	2.167***	2.246***	2.197***	2.203***
	(0.746)	(0.748)	(0.756)	(0.829)	(0.819)	(0.838)
Finance	0.095***	0.116***	0.109***	0.107***	0.136***	0.131***
	(0.031)	(0.032)	(0.032)	(0.033)	(0.033)	(0.033)
Government	-0.022	-0.024	-0.021	-0.021	-0.028	-0.021
	(0.020)	(0.021)	(0.021)	(0.020)	(0.021)	(0.020)
Unemployment	-0.002	-0.004	-0.003	-0.001	-0.003	-0.003
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
TaxK	-0.024***	-0.025***	-0.024***	-0.025***	-0.025***	-0.025***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
TaxL	0.016***	0.017***	0.017***	0.015***	0.015***	0.015***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
R^2	0.853	0.853	0.854	0.860	0.862	0.863
Observations	1530	1530	1530	1377	1377	1377

Notes: Variable description is given in Table 1. Innovation by entrants is a count of innovation that restricts to patents whose assignee first patented less than 5 years ago. Other patents enter in the count of Innovation by incumbents. Both these measures of innovation are taken in log and lagged by two years. Panel data OLS regressions with state and year fixed effects. Time span for innovation: 1980-2009 (columns 1 to 3) and 1980-2006 (columns 4 to 6). Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B6: Top 1% income share and innovation for single state inventors

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.020*	0.040***	0.025**	0.023**	0.020***	0.020***
	(0.012)	(0.009)	(0.010)	(0.011)	(0.005)	(0.004)
Gdppc	0.089**	0.059	0.077*	0.085*	0.077*	0.084*
	(0.043)	(0.045)	(0.044)	(0.044)	(0.044)	(0.044)
Popgrowth	0.862	1.026	0.906	0.898	0.850	0.848
	(0.653)	(0.706)	(0.651)	(0.652)	(0.705)	(0.685)
Finance	0.076**	0.104***	0.082**	0.081**	0.094***	0.094***
	(0.035)	(0.036)	(0.035)	(0.035)	(0.036)	(0.036)
Government	-0.019	-0.020*	-0.019*	-0.019*	-0.016	-0.018
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Unemployment	-0.006**	-0.006**	-0.006**	-0.006**	-0.006*	-0.006*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
TaxK	-0.024***	-0.025***	-0.024***	-0.024***	-0.024***	-0.025***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)
TaxL	0.012***	0.011***	0.013***	0.013***	0.010***	0.010***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
R^2	0.890	0.895	0.890	0.890	0.895	0.895
Observations	1734	1581	1734	1734	1581	1581

Notes: This table shows similar results as the one from Table 6 but patents from inventors that have changed its state of residence over the period are removed. All the innovation measures as well as the dependent variable are taken in log. Time span for innovation: 1976-2009 (columns 1, 3 and 4) and 1976-2008 (columns 2, 5 and 6). Variable description is given in Table 1. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B7: Regression of innovation on top 1% income share using instrument based on Appropriation Committee composition in the Senate. Single state inventors

Dependent variable	Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.213*	0.167**	0.157**	0.201**	0.127**	0.117**
	(0.111)	(0.067)	(0.069)	(0.093)	(0.051)	(0.048)
Gdppc	-0.093	-0.060	-0.104	-0.115	-0.011	0.048
	(0.116)	(0.084)	(0.103)	(0.114)	(0.082)	(0.059)
Popgrowth	1.897**	1.689*	2.017**	2.236**	0.624	0.369
	(0.944)	(0.952)	(0.899)	(1.049)	(0.938)	(0.729)
Finance	0.178**	0.196***	0.160***	0.184***	0.199***	0.192***
	(0.072)	(0.060)	(0.052)	(0.063)	(0.064)	(0.060)
Government	-0.098***	-0.079***	-0.087***	-0.103***	-0.036	-0.041*
	(0.025)	(0.024)	(0.024)	(0.027)	(0.026)	(0.025)
Unemployment	-0.013**	-0.010**	-0.011**	-0.011**	-0.012***	-0.014***
	(0.006)	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)
TaxK	-0.043***	-0.040***	-0.039***	-0.039***	-0.037***	-0.039***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
TaxL	0.021**	0.016**	0.019***	0.021***	0.013*	0.016**
	(0.009)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Highways	0.202	0.359	0.326	0.435	0.376	0.683
	(0.401)	(0.401)	(0.410)	(0.448)	(0.431)	(0.436)
Military	-0.005	-0.005	-0.007	-0.007	-0.007	-0.003
	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
R^2	0.904	0.913	0.876	0.862	0.864	0.870
F-stat on the excluded instruments	11.5	18.2	26.0	15.1	16.2	14.5
Observations	1700	1550	1700	1700	1550	1550

Notes: This table shows similar results as the one from Table 14 but patents from inventors that have changed its state of residence over the period are removed. Time span for innovation: 1976-2009 (columns 1, 3 and 4) and 1976-2006 (columns 2, 5 and 6). Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B8: 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)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation	0.192**	0.077*	0.031	-0.018	-0.048	0.072*
	(0.078)	(0.046)	(0.034)	(0.028)	(0.036)	(0.037)
Gdppc	-0.098	-0.035	-0.031	0.002	-0.002	0.062
	(0.094)	(0.054)	(0.046)	(0.035)	(0.045)	(0.047)
popgrowth	1.691*	0.812	0.442	-0.288	-0.492*	0.607
	(0.997)	(0.587)	(0.451)	(0.204)	(0.285)	(0.375)
Finance	0.226***	0.111***	0.042	-0.003	-0.049	0.084**
	(0.070)	(0.041)	(0.032)	(0.024)	(0.032)	(0.033)
Government	-0.077***	-0.046***	-0.022*	-0.003	0.014	-0.037***
	(0.023)	(0.014)	(0.012)	(0.010)	(0.012)	(0.010)
Unemployment	-0.013**	-0.004	-0.000	-0.000	0.003	-0.005**
	(0.005)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
TaxK	-0.025***	-0.010***	-0.000	-0.004***	-0.001	-0.011***
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
TaxL	0.014***	0.005**	-0.001	0.002	-0.001	0.009***
	(0.004)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Highways	0.386	0.153	-0.035	-0.014	-0.110	0.144
	(0.470)	(0.263)	(0.176)	(0.128)	(0.168)	(0.242)
Military	-0.006	-0.008*	-0.010**	-0.009**	-0.009**	-0.008**
	(0.007)	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)
R^2	0.872	0.806	0.430	0.864	0.712	0.934
F-stat on excluded instruments	14.2	14.2	14.2	14.2	14.2	14.2
Observations	1550	1550	1550	1550	1550	1550

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by two years. The dependent variables are also taken in log. 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 endogenous variable is set to 3 years. Time span for innovation: 1976-2009 for columns 1, 3 and 4 and 1976-2006 for columns 2, 5 and 6. DC is removed from the sample because it has no senators. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B9: Top 1% income share and innovation at different lags - IV results

Dependent variable	Log of Top 1% Income Share				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5
Lag of innovation	2 years	3 years	4 years	5 years	6 years
Innovation at $t-2$	0.253**				
	(0.112)				
Innovation at $t-3$		0.180**			
		(0.083)			
Innovation at $t-4$			0.176**		
			(0.078)		
Innovation at $t-5$				0.071	
				(0.065)	
Innovation at $t-6$					0.057
					(0.065)
Gdppc	-0.190	-0.097	-0.077	0.033	0.059
	(0.163)	(0.133)	(0.121)	(0.102)	(0.097)
popgrowth	2.300*	3.005**	3.072**	2.806***	2.641***
	(1.263)	(1.289)	(1.194)	(1.061)	(1.006)
Finance	0.292***	0.220***	0.187***	0.115**	0.097**
	(0.108)	(0.078)	(0.064)	(0.048)	(0.042)
Government	-0.088**	-0.088**	-0.083**	-0.110***	-0.116***
	(0.036)	(0.035)	(0.036)	(0.033)	(0.032)
Unemployment	-0.022**	-0.014**	-0.012**	-0.008*	-0.007*
	(0.009)	(0.006)	(0.005)	(0.004)	(0.004)
TaxK	-0.034***	-0.032***	-0.031***	-0.031***	-0.031***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
TaxL	0.022**	0.022***	0.021**	0.016**	0.014**
	(0.009)	(0.009)	(0.008)	(0.007)	(0.007)
Highways	2.302***	1.889***	1.883***	1.537***	1.503***
	(0.648)	(0.476)	(0.474)	(0.430)	(0.433)
Military	-0.009	-0.008	-0.007	-0.004	-0.003
	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)
R^2	0.856	0.868	0.843	0.842	0.839
F-stat on excluded instruments	17.8	17.8	17.8	17.8	17.8
Observations	1450	1450	1450	1450	1450

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by 2 to 6 years. 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 endogenous variable is set to 3 years. Time span for innovation: 1979-2006. DC is removed from the sample because it has no senators. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B10: Regression of innovation on Top 1% income share using only the spillover instrument

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Patents	Cit5	Claims	Generality	Top5	Top1
Innovation	0.183***	0.153***	0.112***	0.113***	0.099***	0.147***
	(0.057)	(0.037)	(0.034)	(0.032)	(0.024)	(0.048)
Gdppc	-0.071	-0.045	-0.038	-0.016	-0.035	-0.071
	(0.085)	(0.069)	(0.077)	(0.070)	(0.070)	(0.079)
Popgrowth	2.385***	2.403**	2.400***	2.184***	2.149**	2.937***
	(0.899)	(0.954)	(0.843)	(0.800)	(0.913)	(1.014)
Finance	0.197***	0.239***	0.157***	0.161***	0.205***	0.266***
	(0.045)	(0.045)	(0.037)	(0.037)	(0.042)	(0.067)
Government	-0.028	-0.020	-0.034*	-0.033	0.001	0.030
	(0.020)	(0.019)	(0.020)	(0.020)	(0.020)	(0.028)
Unemployment	-0.011***	-0.009**	-0.008**	-0.008**	-0.008**	-0.006
	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
TaxK	-0.026***	-0.025***	-0.024***	-0.025***	-0.026***	-0.024***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
TaxL	0.021***	0.017***	0.018***	0.019***	0.015***	0.017***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
R^2	0.828	0.843	0.837	0.842	0.826	0.743
F-stat on excluded instruments	29.3	36.0	52.3	73.9	46.2	17.0
Observations	1530	1377	1530	1530	1377	1377

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by 2 years. Panel data IV 2SLS regressions with state and year fixed effects. Innovation is instrumented by a measure of spillover as described in section 6.3. The lags between the instruments and the endogenous variable is set to 1 year. Control for spatial correlation involves adding two additional controls for demand shocks as explained in subsection 6.3. Time span for innovation: 1981-2006. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B11: Robustness 2: financial sector and natural resources - IV results

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation	0.197***	0.265***	0.196**	0.200*	0.157**	0.195**
	(0.009)	(0.006)	(0.016)	(0.051)	(0.024)	(0.015)
Gdppc	-0.258**	-0.180	-0.097	-0.099	0.032	-0.093
	(0.011)	(0.112)	(0.311)	(0.322)	(0.696)	(0.320)
popgrowth	1.637	1.778	1.635	1.725*	1.698*	1.694*
	(0.107)	(0.117)	(0.104)	(0.097)	(0.051)	(0.090)
Finance	0.151**	0.221***	0.231***	0.230***	0.226***	0.228***
	(0.016)	(0.001)	(0.001)	(0.004)	(0.000)	(0.001)
Government	-0.031	-0.066***	-0.079***	-0.079***	-0.069***	-0.078***
	(0.149)	(0.006)	(0.001)	(0.001)	(0.001)	(0.001)
Unemployment	-0.018***	-0.020***	-0.013**	-0.013***	-0.010**	-0.013**
	(0.001)	(0.002)	(0.014)	(0.008)	(0.036)	(0.013)
TaxK	-0.023***	-0.022***	-0.025***	-0.025***	-0.020***	-0.025***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
TaxL	0.012***	0.013**	0.014***	0.014***	0.014***	0.014***
	(0.001)	(0.016)	(0.001)	(0.001)	(0.000)	(0.001)
Highways	-0.284	0.169	0.410	0.414	0.577	0.378
	(0.526)	(0.749)	(0.395)	(0.431)	(0.133)	(0.423)
Military	0.006	-0.007	-0.006	-0.006	-0.004	-0.006
	(0.362)	(0.386)	(0.436)	(0.449)	(0.586)	(0.435)
RemunFinance	0.396***					
	(0.000)					
EFD				-0.177		
				(0.730)		
Mining+Oil					0.041***	
					(0.000)	
R^2	0.885	0.834	0.873	0.870	0.894	0.874
F-stat on excluded instruments	14.1	10.8	13.5	10.5	15.2	13.8
Observations	1550	1426	1550	1550	1550	1550

Notes: Variable *Mining+oil* measure the share of oil related and natural resources extraction activities in GDP, variable *RemunFinance* measures the compensation per employee in the financial sector and variable *EFD* measures the financial dependence of innovation. Other variables description is given in Table 1. Innovation is taken in log and lagged by two years. Column 1 controls for average compensation in the financial sector, column 2 drops NY, CT, DE and MA (the state with the largest financial sectors), column 3 removes finance-related patents, column 4 controls for financial dependence in the state as explained in section 6.1, column 5 controls for the size of oil and mining sectors and column 6 removes oil-related patents from the count of citations. Time Span for innovation: 1976-2008. 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 endogenous variable is set to 3 years. Time span for innovation: 1979-2006. DC is removed from the sample because it has no senators. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B12: Robustness 3: controlling for industry composition - OLS results

Dependent variable	Log of Top 1% Income Share				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation	0.036***	0.050***	0.049***	0.050***	0.048***
	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)
Gdppc	0.075*	0.049	0.051	0.058	0.045
	(0.045)	(0.044)	(0.044)	(0.043)	(0.044)
popgrowth	1.027	1.047	1.041	1.197*	1.039
	(0.704)	(0.706)	(0.706)	(0.707)	(0.704)
Finance	0.091**	0.113***	0.112***	0.132***	0.116***
	(0.036)	(0.036)	(0.036)	(0.044)	(0.036)
Government	-0.022*	-0.019*	-0.020*	-0.023**	-0.020*
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Unemployment	-0.005*	-0.007**	-0.007**	-0.007**	-0.007**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
TaxK	-0.024***	-0.025***	-0.025***	-0.024***	-0.025***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
TaxL	0.011***	0.011***	0.011***	0.011***	0.011***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Size of Sector:					
Computer and Electronic				0.429	
				(0.485)	
Chemistry				-0.608***	
				(0.184)	
Electrical Component				4.380**	
				(2.036)	
R^2	0.895	0.897	0.897	0.898	0.897
Observations	1581	1581	1581	1578	1581

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by 2 years. 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. Panel data OLS regressions with state and year fixed effects. Time span for innovation: 1979-2006. Autocorrelation and heteroskedasticity robust standard errors using the Newey-West variance estimator are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table B13: ROBUSTNESS 4: CONTROLLING FOR AGGLOMERATION EFFECT - OLS AND IV RESULTS.

Dependent variable	Log of Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Measure of innovation	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation	0.047***	0.048***	0.047***	0.200***	0.199***	0.200***
	(0.008)	(0.008)	(0.008)	(0.077)	(0.077)	(0.077)
Gdppc	0.036	0.035	0.033	-0.114	-0.113	-0.116
	(0.044)	(0.045)	(0.044)	(0.094)	(0.094)	(0.094)
Popgrowth	1.124	1.122	1.122	1.749*	1.745*	1.749*
	(0.719)	(0.720)	(0.719)	(1.027)	(1.025)	(1.025)
Finance	0.104***	0.108***	0.107***	0.227***	0.229***	0.229***
	(0.035)	(0.035)	(0.035)	(0.069)	(0.069)	(0.069)
Government	-0.055***	-0.054***	-0.054***	-0.071***	-0.072***	-0.072***
	(0.012)	(0.012)	(0.012)	(0.023)	(0.023)	(0.023)
Unemployment	-0.009***	-0.008***	-0.009***	-0.015***	-0.015***	-0.015***
	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)
TaxK	-0.023***	-0.023***	-0.023***	-0.024***	-0.024***	-0.024***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
TaxL	0.013***	0.013***	0.013***	0.016***	0.016***	0.016***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Agglomeration	0.185***	0.174***	0.178***	0.153***	0.144***	0.145***
	(0.019)	(0.018)	(0.018)	(0.021)	(0.021)	(0.022)
Highways				0.207	0.211	0.207
				(0.483)	(0.479)	(0.483)
Military				-0.005	-0.006	-0.006
				(0.007)	(0.007)	(0.007)
R^2	0.903	0.903	0.903	0.872	0.873	0.872
F-stat on excluded instruments				14.3	14.2	14.2
Observations	1581	1581	1581	1550	1550	1550

Notes: Variable description is given in Table 1. Innovation is taken in log and lagged by 2 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 1. 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 endogenous 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.

C Details on the calibration

In this section we explain how the 6 moments in the data determine the 6 parameters of the model (θ_I, θ_E, ϕ, L, η_H, η_L). For convenience, we introduce the following notations: $t = M_1$denotes the top 1% income share, $R = M_2$ denotes the ratio of entrant to incumbent innovations, $e = M_3$ denotes the elasticity of the top 1% share with respect to innovation, $\widetilde{\eta} = M_4$ denotes the average mark-up, $E = M_5$ denotes the entrants' employment share and $g = M_6$ denotes the growth rate. By definition the innovation ratio obeys $x_E = Rx_I$,the growth rate of the economy is given by:

$$(\phi x_I + x_E) \ln \eta_H = g, \tag{36}$$

and the average mark-up by:

$$\mu \eta_H + (1 - \mu) \eta_L = \widetilde{\eta}. \tag{37}$$

Provided that $\mu / (1 + L) < 1/100 < 1/(1 + L)$ (which ex-post ends up being the relevant range), (12) implies that the top 1% share obeys:

$$\mu \left(\frac{1}{\eta_L} - \frac{1}{\eta_H} \right) + \frac{1 + L}{100} \left(1 - \frac{1}{\eta_L} \right) = t, \tag{38}$$

and, using (13), the semi-elasticity of the top 1% share with respect to innovation is given by:

$$\mu \left(\frac{1}{\eta_L} - \frac{1}{\eta_H} \right) = et. \tag{39}$$

The semi-elasticity of innovation with respect to the top 1 % share depends on the innovation rate μ and on the extra-profit share made by innovating entrepreneurs, namely $1/\eta_L - 1/\eta_H$.

Finally, using that labor costs equal revenues divided by the mark-up, we obtain that the entrant employment share is given by:

$$E = \frac{x_E \frac{1}{\eta_H}}{\mu \frac{1}{\eta_H} + (1 - \mu) \frac{1}{\eta_L}}. \tag{40}$$

In the above equations, we have that $\mu = x_I + x_E$ (since $z = 0$), x_I is given by (9) and, using (11), x_E is given by:

$$x_E = \frac{1 - \frac{1}{\eta_H} - \frac{1}{L} \frac{1}{\eta_L} + \frac{1}{L} \left(\frac{1}{\eta_L} - \frac{1}{\eta_H} \right)^2 \frac{1}{\theta_I}}{\theta_E - \frac{1}{L} \left(\frac{1}{\eta_L} - \frac{1}{\eta_H} \right)}. \tag{41}$$

Assume that one knows μ, η_L and η_H. Then one obtains $x_E = \frac{R}{1+R}\mu$ and $x_I = \frac{1}{1+R}\mu$. Given the innovation rates and the innovation step-size, the rate of productive incumbent innovations, ϕ, is identified by the measured growth rate in the economy, following (36). Given the extra profits made by innovative entrepreneurs ($\mu\left(\frac{1}{\eta_L} - \frac{1}{\eta_H}\right)$), the top 1% share in (38) determines the number of non-innovative entrepreneurs that are in the top 1% share and thereby identifies the ratio of workers to entrepreneurs L. Given x_I, x_E, η_L and η_H, the R&D parameters θ_I and then θ_E are directly given by (9) and (41).

In return, the innovation rate μ and the mark-ups η_L and η_H are determined solely by 3 equations. Equation (37) gives the average mark-up. Equation (39) gives a relationship between the innovation rate and the difference in inverse mark-ups. (40) can be rewritten as

$$\frac{R}{1+R}\mu\frac{1}{\eta_H} = E\left(\mu\frac{1}{\eta_H} + (1-\mu)\frac{1}{\eta_L}\right). \tag{42}$$

Given the harmonic average of mark-ups, a high innovation rate increases the entrant employment share but a high innovation step reduces it.

In fact, we can go a bit further, combining (39) and (37), we get:

$$\eta_L = \widetilde{\eta}/\left(et\eta_H + 1\right). \tag{43}$$

Using this equation and (39), in (42), we get:

$$\left(\frac{et\eta_H + 1}{\widetilde{\eta}} - et\right)\left(\frac{\eta_H\left(et\eta_H + 1\right)}{\widetilde{\eta}} - 1\right) = \frac{R}{1+R}\frac{et}{E} \tag{44}$$

The left-hand side is an increasing function of η_H for $\eta_H \geq \widetilde{\eta}$ and for $\eta_H = \widetilde{\eta}$, it is equal to et, which is lower than the righ-hand side since $\frac{R}{1+R} > E$. Therefore this equation identifies η_H uniquely. And everything else equal, a higher E and a higher R lead to a lower η_H. It is then easy to obtain η_L through (43), where a higher $\widetilde{\eta}$ and lower et and η_H lead to a higher η_L. Finally given η_L and η_H, the innovation rate is determined by (39): a higher semi-elasticity of innovation et, leads to a higher innovation rate e. All parameters can then be indentified as explained above and it can be checked that we are indeed in the case where $\mu/\left(1+L\right) < 1/100 < 1/\left(1+L\right)$.

Calibrating the shocks for the regression. We assume that the shocks to innovation are normally distributed with $\varepsilon_{\theta,i,t} \sim N\left(0.72^2/2, 0.72^2\right)$ and $\varepsilon_{\theta,i} \sim N\left(0.62^2/2, 0.62^2\right)$ (this implies that the R&D shock to $x_{I,i,t} = \frac{1}{\theta_{I,i,t}}\left(\frac{1}{\eta_{L,i,t}} - \frac{1}{\eta_H}\right)$ has mean 0). The shocks to non-innovator mark-ups are uniformly distributed with $\varepsilon_{\eta,i,t} \sim U\left(-0.6g, 0.6g\right)$ and $\varepsilon_{\eta,i,t} \sim U\left(-0.4g, 0.4g\right)$ where $g = \min\left(\eta_L - 1, \eta_H - 1\right)$ (so that we always have $\eta_H > \eta_{L,i,t} > 1$).

With the different shocks it is possible that $x_{I,i,t}$ or $x_{E,i,t}$ are greater than 1 which, within the model, makes no sense, therefore we censor the two variables at 1. The "measurement error shocks" on top income inequality are normally distributed with $\varepsilon_{\delta,t} \sim N\left(-0.325^2/2, 0.325^2\right)$, $\varepsilon_{\delta,i,t} \sim N\left(-0.07^2/2, 0.07^2\right)$ and $\varepsilon_{\delta,i} = 0$ in the specific regressions we report on. The measurement error to innovation $\varepsilon_{\mu,i,t}$ is also normally distributed with $\varepsilon_{\mu,i,t} \sim N\left(-0.47^2/2, 0.47^2\right)$. The standard deviations are chosen so that the simulated data (an average of 500 draws) are in line with the actual data regarding the empirical moments reported in Table C1 below.[60]

Table C1: L

Moment	Empirics	Simulation	Affected by the s.d. of:
s.d. of the state fixed effects	0.173	0.163	$\varepsilon_{\delta,i}$, $\varepsilon_{\theta,i}$ and $\varepsilon_{\eta,i}$
s.d. of the year fixed effects	0.325	0.319	$\varepsilon_{\delta,t}$
s.d. of log Cit5	1.348	1.107	$\varepsilon_{\theta,i,t}$, $\varepsilon_{\theta,i}$, $\varepsilon_{\eta,i,t}$, $\varepsilon_{\eta,i}$ and $\varepsilon_{\mu.i,t}$
s.d.of log Cit5 controlling for state fixed effects	0.744	0.889	$\varepsilon_{\theta,i,t}$, $\varepsilon_{\eta,i,t}$ and $\varepsilon_{\mu.i,t}$
s.d. of predicted log Cit5	1.057	0.959	$\varepsilon_{\theta,i,t}$, $\varepsilon_{\theta,i}$, and $\varepsilon_{\eta,i}$
s.d.of predicted log Cit5 controlling for state fixed effects	0.702	0.697	$\varepsilon_{\theta,i,t}$
s.d. of top income inequality	0.046	0.056	all except $\varepsilon_{\mu,i,t}$
OLS coefficient	0.049	0.051	$\nearrow$ with the s.d. of $\varepsilon_{\theta,i,t}$
			$\swarrow$ with the s.d. of $\varepsilon_{\eta,i,t}$ and $\varepsilon_{\mu,i,t}$

Notes: Standard deviation of various variables in the data and on average in 500 draws of our simulated data.

This shows that "realistic" shocks in terms of the deviations observed in the data can reproduce the gap between the OLS and IV coefficient. The major deviation in the table above is that in our simulations, the standard deviation of the (unpredicted) innovation variable (log Cit5) is too large when one controls for state fixed effects and too small when one does not. This could be corrected for instance by introducing state-specific measurement errors on innovation.

[60]We do not try to identify what the standard deviations of the shocks are and therefore did not try to choose standard deviations so as to minimize the distance between empirical moments and simulated moments here.

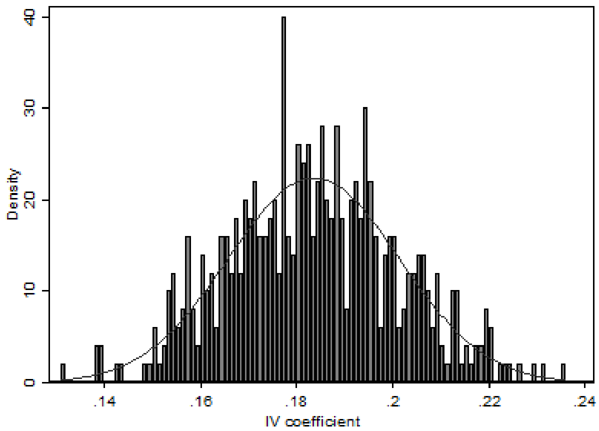


Figure 8: Distribution of the IV coefficients

Notes: This figure plots the whole distribution of the IV coefficient of innovation obtained with 1000 draws. Average value is 0.183 and standard deviation is 0.0178.

D Computing top income shares at the CZ level

To compute top income shares at the CZ level, we need to estimate what is the expected income for individuals whose income is censored. To do that we expend on the methodology of Clemens et al. (2017) (see also Armour et al., 2016). The census reports separately individuals' labor income, capital income and business income and each income source is censored separately. We assume that above a certain total income level $\underline{x}$, income is Pareto distributed.

Denote by $\bar{l}$, $\bar{c}$ and $\bar{b}$ the levels above which the census data are censored. Then for an individual i, denote by l_i her labor income, c_i her capital income, and b_i her business income *as reported in the data* with $l_i = \bar{l}$ if labor income is censored and similarly $b_i = \bar{b}$ or $c_i = \bar{c}$ if another source of income is censored. Denote x_i the true total income of individual i. Then there are two cases. First, if $l_i < \bar{l}$, $c_i < \bar{c}$ and $b_i < \bar{b}$, then we know that her total income is $x_i = l_i + b_i + c_i$. Conditional on having $x_i > \underline{x}$, the conditional probability density function of observing an income x_i is given by $\alpha \underline{x}^{\alpha} / x_i^{\alpha+1}$ where α is the shape parameter of the Pareto distribution of income. Denote by $\mathcal{N}_{unc}$ the corresponding set of observations for which no information is censored and where $x_i > \underline{x}$, and by N_{unc} the cardinal of that set.

On the other hand, if one observation or more are censored, then we only know that her total income $x_i \geq \overline{x}_i \equiv l_i + b_i + c_i$. Conditional on having $\overline{x}_i \geq \underline{x}$, then the probability of observing $x_i \geq \overline{x}_i$ is given by $(\underline{x}/\overline{x}_i)^{\alpha}$. Denote by $\mathcal{N}_{cens}$ the corresponding set of observations for which at least one source of income is censored and with $\overline{x}_i \geq \underline{x}$ (we choose $\underline{x}$ low enough so that this is always the case when an observation is censored).

We can then write the likelihood function as

$$P = \prod_{i \in \mathcal{N}_{unc}} \alpha \left(\frac{\underline{x}}{x_i} \right)^{\alpha} \frac{1}{x_i} \prod_{i \in \mathcal{N}_{cens}} \left(\frac{\underline{x}}{\overline{x}_i} \right)^{\alpha}.$$

The resulting maximum log-likelihood estimate is given by

$$\frac{1}{\widehat{\alpha}} = \frac{1}{N_{unc}} \left(\sum_{i \in \mathcal{N}_{unc}} \ln \left(\frac{x_i}{\underline{x}} \right) + \sum_{i \in \mathcal{N}_{cens}} \ln \left(\frac{\overline{x}_i}{\underline{x}} \right) \right).$$

We can then computing top income shares by assuming that for any observation $i \in \mathcal{N}_{cens}$, we have that $x_i = \frac{\widehat{\alpha}}{\widehat{\alpha}-1} \overline{x}_i$.