Innovation and Inequality: World Evidence

Benos, Nikos and Tsiachtsiras, Georgios

University of Ioannina

27 September 2018

Innovation and Inequality: World Evidence

Nikos Benos Georgios Tsiachtsiras

Abstract

In this paper we use country panel data to explore the effect of innovation on top income inequality. We construct a novel dataset of patents by combining patents from USPTO and EPO to test the effect of innovation on income inequality. We demonstrate that innovation has a strong positive correlation with top income shares. Also, we find weak evidence that innovation has a negative effect on overall income inequality. We support our findings by using instrumental variables to tackle endogeneity. In addition our IV analysis shows that the effect of innovation on top income shares remains significant for 3 years. Finally, we show that innovation has a less strong effect on top income inequality when we include defensive patents in the analysis.

JEL classification: D63, O30, O31, O33, O34, O40, O47

Keywords: top income inequality, overall inequality, innovation, citations, defensive patents.

Contact Information: Benos: University of Ioannina, Department of Economics, email: nbenos@cc.uoi.gr (corresponding author). Tsiachtsiras: Universitat de Barcelona, School of Economics, email: ec002543@gmail.com.

Acknowledgements: We would like to thank Anastasia Litina for very helpful comments and discussions. In addition we would like to thank Evangelos Dioikitopoulos, Dimitrios Dadakas and Nikos Tsakiris as well as the participants of the 4th International Conference on Applied Theory, Macro and Empirical Finance (University of Macedonia, 2018) for their comments and suggestions. We would also like to show our gratitude to the Science, Technology and Innovation Microdatalab of OECD for providing us the data to carry out this research. Any remaining errors are ours.

1 Introduction

Many recent studies show that there is a steady rise in income inequality. According to them one factor behind this phenomenon is innovation. Although the literature confirms that innovation plays a crucial role in the evolution of income inequality, there is no yet a definitive answer on whether this effect is positive or negative. Also, the direction of causality between inequality and innovation is not clear yet.

Figure 1 illustrates the trends of citations and top one percent income share. After 1990 we see an increase in both innovation and top income inequality until 2000 where innovation reaches at its peak. Afterwards, both variables remain high until 2005 after which they start to fall dramatically. We conclude that they have parallel trends for many years. In Figure 2 we present a scatter plot of the log differences of citations years and top 1% income share. The linear fit reveals a positive correlation between innovation and top 1%. This evidence and the work of Aghion et al. (2018) about top income inequality in USA inspired us to test the effect of innovation on top income inequality at the country level using a world sample.

In this paper we argue that innovation is positively associated with top income shares. We compose a novel dataset of patents by including patents from EPO and USPTO. More details about the construction exist in the Appendix. We use OLS regressions with country and year fixed effects to explore this relationship among different countries over time. We also apply IV regressions to check the robustness of our basic results and confirm that innovation boosts income inequality. Our main contributions in the literature are three. First, we contribute to the literature on inequality and growth by using innovation as a channel linking the two (Aghion and Howitt, 1992). Second, we enrich the literature about innovation and income inequality by including in our analysis different measures of both innovation and income inequality (Aghion at el., 2018). Last, we analyze the influence of the defensive patents on this relationship (Abrams et al., 2013).

The rest of the paper is organized as follows. Section 2 explores the relations between innovation and income distribution. Section 3 presents the data. Section 4 shows the empirical strategy. Section 5 discusses the results and Section 6 concludes.

2 Theoretical framework

According to many studies innovation has a positive effect on income inequality. First, there is the productivity effect which boosts wages of employees who work in innovative firms (Lee, 2011). These firms are able to develop new products and as a result new jobs are created (Breau et al., 2014). The new jobs require advanced technologies suitable only for high skilled employees and this impact shows up in their salaries (Lee, 2011 and Breau et al., 2014). Also, innovative regions lure highly skilled and highly paid workers (Lee, 2011).

Innovation may have different results in countries with dissimilar institutions. For example, Scandinavian countries prefer egalitarian societies (Acemoglu et al., 2012). Also, in contrast with many European countries, the flexible US markets allow high skilled individuals to enter innovative sectors (Lee and Pose, 2012).

The effect of innovation is stronger on top income shares than the rest of the income distribution (Aghion at el., 2018). According to them, innovation from both incumbents and entrants increases top income inequality. The difference between incumbents and entrants is that incumbents erect barriers. The barriers discourage new entrants and boost top income inequality. Also, the authors propose an additional channel through which innovation affects top income inequality. This channel is capital gains. The source of capital gains is the award for the innovative companies (mark-up). They indicate that through mark-up the companies have managed to increase their profits during the past forty years. Entrepreneurs and CEOs earn the bigger share of profits.

There are empirical findings, which confirm all above arguments. Lee (2011) uses data from the European Community Household Panel for the period 1995-2001 and finds that innovation has a positive effect on income inequality. The results are similar for the Canadian cities (Breau et al., 2014), while Aghion et al. (2018) focus on top income shares for US states. Aghion et al. (2018) conclude that innovation drives inequality and not the opposite by applying IV regressions.

However many studies indicate that innovation is the key to reduce income inequality. There are many arguments, which support this finding. Innovation creates knowledge spillovers (Aghion at el., 2018), which can benefit individuals with fewer skills (Lee and

Pose, 2012). These individuals can learn from their high skilled partners and augment their productivity (Lee, 2011). They then manage to increase their salaries and income inequality falls.

Apart from the spillovers effect, economic growth may reduce income inequalities (Antonelli and Gehringer, 2013). According to them economic growth increases wages of all individuals in the economy. They state that the strong price competition among companies could decrease the accumulation of rents. Economic growth reduces interest rates and this in turn causes a fall in capital gains. They conclude that in a Schumpeterian framework with fast rate of technological change the reduction of income inequality is possible. However, they recognize that market imperfections have negative consequences on the correct allocation of resources in favor of the richest people. Aghion et al. (2018) use also in their paper a Schumpeterian framework and panel data from USA. Even though they use an economy with fast rate of technological change they find significant evidence that innovation has an effect on income inequality after many years. It seems that market imperfections and barriers from incumbents help to maintain the effect of innovation at least in the short run. Antonelli and Gehringer (2013) base their findings on a big sample of European countries, USA, Canada, China, Korea and India.

Also inequality can affect innovation. A decrease in inequality may trigger an increase in the number of customers who can buy new products (Hatipoglu, 2012). The change in inequality can affect the inventors' expected profits and their decisions about R&D investments. In addition, this article strengthens our suspicions about the potential endogeneity problem between innovation and income inequality. We try to solve this problem in a next section.

3 Data

The data on pre-tax income share owned by the top 10%, 1%, 0.5% and 0.1% of income earners in our country panel analysis are drawn from the World Wealth and Income Database (Alvaredo et al., 2017). These data are available for some countries from 1870 to 2016 but we focus on the period after 1960[1]. We subtract the top 1%

[1] The time series data for the rest of the control variables starts in 1960.

income share from the top 10% of income share and then we divide it by nine to create the average top income share (Aghion et al, 2018). In addition, we use the Theil index. We extract the Theil index from the University of Texas Database which covers the time period 1963-2015 for 151 countries (Galbraith et al., 2013). We have chosen the Standardized World Income Inequality Database (Frederick Solt, 2016) for the Gini index. The Standardized World Income Inequality Database provides us with 100 equivalent Gini indexes for the pre-tax income and everyone of these has a different standard deviation. We include in our analysis the Gini index with the smallest standard deviation. Again we use data only after 1960.

We apply many measures of innovation. The quantity measures of innovation come from WIPO like in the papers of Hatipoglu (2012) and Antonelli and Gehringer (2013). The first one is the total number of patents granted (direct and PCT national phase entries) and count by filing office. The two other measures, which we use as proxies for the number of patents, are the number of residents' applications and the number of non residents' applications. The number of patents is a crude measure of innovation because a patent with a great contribution to the literature and a patent with small contribution receive the same weight (Aghion at el., 2018).

This is why we apply also quality measures of innovation like citations and the family size of patents. The Science, Technology and Innovation Microdatalab of OECD has provided us with the databases containing quality measures of innovation. The basic Database is the OECD Patent Quality Indicators Database which has the quality measures of innovation: citations on a 5-year window, citations on a 7-year window, top 1% most cited patents and the family size of each patent. In contrast with Aghion et al. (2018) our measures of citations do not suffer much from truncation bias because the citations are included in the patent document within the first two years since application (Squicciarini et al., 2013). We use citations as a measure of innovation because novel innovations will have larger mark-ups due to their originality (Abrams et al., 2013). In addition, these innovations will generate spillovers for subsequent innovations (Abrams et al., 2013). The family size is represented by the number of patent offices at which a given invention has been protected. The most valuable patents are being protected from many different patent offices (Squicciarini et al., 2013). We provide descriptive details about the construction of the databases in the Appendix.

The Patent Quality Indicators Database includes a variable called grant lag. If the value of this variable is high then the patent was granted very fast. In contrast if it is missing value then the patent was not granted. This fact allows us to construct two separate databases, one with just the patents granted and one with the total number of patents. In the main regressions we use only patents that have been granted. However we use the second database as a robustness check and to test the hypothesis about defensive patents (Aghion at el., 2018). This means that companies make strategic patenting to protect their most valuable patents (Abrams et al., 2013). Aghion et al. (2018) use citations as a measure of innovation to address this problem. It is logical that the effect of innovation on income inequality is smaller when we use the full sample of patents[2]. Our purpose is to test this hypothesis.

We extract the rest of the control variables from the World Bank database. We have used the domestic credit (provided by financial sector as a percentage of GDP) to control for the financial sector influence on inequality. The financial sector usually helps the inventors to innovate and increase their salaries (Aghion at el., 2018). A big share of the employees (almost 27%) who belong in the top 0.1% income share in the United States work in the financial sector or use financial services (Szymborska, 2016). The second variable is the general government final consumption expenditure (% of GDP) in order to control for the government size in each country. Empirical studies find that government size has a negative effect on capital income inequality and more specifically on the top 1% income share (Luo et al., 2017). Next we include in our analysis GDP per capita (constant 2010 US$). It has been found that GDP per capita has a positive effect on the overall Gini index and on the highest quintile income shares (Barro, 2008). We also control for the business cycle by using the unemployment rate (Aghion at el., 2018) and also include population growth. We end up with an unbalanced panel of 32 countries over the time period 1960-2015.

4 Estimation Methodology

Our estimation method is similar with the estimation method of Aghion et al., (2018). First we aggregate the number of citations and the family size of each patent at the

[2] We include patent applications and patents that have been granted in the full sample. According to Aghion et al. (2018) defensive applications receive fewer citations than the novel applications.

country level. We standardize citations and family size by dividing them with the number of patents granted in the Patent Quality Indicators database. Next we divide the quantity and quality measures of innovation with population. After that we take the log of our measures of innovation, inequality and GDP per capita. We estimate the following equation:

$$log(y_{it}) = A+B_i+B_t+b_1 log(innov_{i(t-1)}) +b_2 X_{it}+\varepsilon_{it}$$

where i stands for country i, t stands for time period t, y_{it} is the measure of inequality (in log), A is the constant, B_i, B_t correspond to country and year fixed effects, $innov_{i(t-1)}$ is innovation in year $t-1$ (also in log) and, X are the other control variables. We use year and country fixed effects to account for permanent cross-country differences in inequality and overall changes in inequality respectively. The advantage by taking both the measure of inequality and the measure of innovation in logs is that b_1 can be interpreted as the elasticity of inequality with respect to innovation. We estimate autocorrelation and heteroskedasticity robust standard errors in all our regressions.

We have decided to take the one year lag of innovation as independent variable. Our base, Patent Quality Indicators, provides us with the application dates of the patents. We include in our analysis patents from both EPO and USPTO. Depalo and Addario (2014) use the EPO patents Database and find that inventors' earnings peak at $t-1$ instead of t. They assume that bureaucracy is responsible for this delay. A second empirical study of Bell (2016), who uses patents from USPTO, confirms the conclusion that inventors' income increases before the application date. As a result, we choose the one year lag for our measures of innovation.

In the second part of our empirical analysis, we try to tackle the endogeneity problem between innovation and income inequality. Our instrument is the "charges for the use of intellectual property, receipts (BoP, current US$)" from International Monetary Fund. Our argument is that countries which possess patents of high quality are going to receive bigger amounts of money for the authorized use of proprietary rights such as patents, trademarks, copyrights, industrial processes and designs including trade secrets and franchises. This is the first reason why we are interested on the receipts and not payments. Intellectual property rights have a positive effect on measures of innovation. Strong protection stimulates innovative activity and increases innovation incentives

(Kanwar and Evensont, 2003). Kanwar and Evensont (2003) find that intellectual property protection has a positive effect on R&D expenditures. Dutta and Sharma (2008) test the effect of intellectual property rights on Indian firms and they find that not only IPR increases R&D expenditures but also facilitates patenting by India in the U.S. Chu et al., (2017) explores the effect of IPR on China. They conclude that IPR has a positive effect on innovation. Branstetter et al. (2005) use U.S. multinational firms' data and find limited evidence that IPR boosts domestic innovation. The literature confirms the positive relationship between IPR and innovation.

Next we examine if this instrument is exogenous to income inequality. Here is the second reason why we use receipts. The "charges for the use of intellectual property, receipts" come from non-residents. At the country level this means that this amount of money enters the domestic market from a foreign country. So, we believe that this instrument correlates directly only with our measures of innovation and it is unlikely to affect other domestic variables. To avoid any suspicions that our variable could potential affect indirectly our measures of inequality we use the lead of the variable as instrument. By using the $t+1$ period for our instrument we believe that the case for it being exogenous with regard to income inequality in period t is even stronger. There is a second reason for using the $t+1$ value of our instrument. The grant lag variable, from the Patent Quality Indicators database, has an average mean 4.5 for the EPO and an average mean 2 approximately for the USPTO. If we combine the two datasets we have an average mean of 3 years. This means that a patent needs 3 years from the application date to be granted. We apply the $t-1$ year to the application date in our model and we use the $t+1$ year of our instrument. These are 2 years and we are very close to the average mean of the combined dataset. To avoid losing more observations from our sample we stop to the one year lead of our instrument. Also it is common that companies sell a product embedding an innovation before the patent has been granted (Aghion at el., 2018).

5 Results

In this section we present the results from both OLS and IV regressions. All the variables are defined in Table 1. First we provide the sample of countries in Table 2 sorted both by the number of patents and top income share. Seven of the most patenting countries belong to the top 15 countries with the biggest top income share. Then we present summary statistics for all variables in Table 3. In Tables 4 and 5 we provide descriptive statistics for the measures of innovation and inequality for two distinctive years. It is clear that there is a significant increase in the means of our measures of inequality from 1990 to 2010. Also the min and the max values increased over these years. We reach the same conclusion also from the table with the innovation measures.

Next, we provide the results from OLS regressions. Table 6 regresses the top 1% income share on our measures of innovation with a 1-year lag. We see that only the citations have a significant and positive effect on top 1% income share as we expect from theory. We have taken the logs for both measures of innovation and top income shares so that we can interpret the coefficient of innovation as elasticity. A 1% increase in the number of citations is associated with a 0.0315% increase of the top 1% income share. In contrast with citations, family size has no effect on the top 1% income share. Two out of three quantity measures of innovation have no effect on the top 1% share but residents' applications appear to have a negative and significant effect. The rest of the variables in column 4 have the expected signs. In Table 7 we use cluster standard errors at the country level. The citations on a five-year window keep the positive effect but at the 10% level of significance. We test the effect of innovation on different top income shares in Table 8. The magnitude of the effect is bigger on the top 0.1% income share. Next, we test the effect of innovation on different measures of inequality in Table 9. It is clear from the table that innovation influences only the top 1% income share. In Table 10 we apply different lags of innovation on top 1% income share. We find evidence that the effect of innovation is significant for six years. This fact implies that the Schumpeterian framework does not work very fast and innovation boosts income inequality in contrast with the findings of Antonelli and Gehringer (2013).

As we said above we have tried to address the endogeneity problem. There are empirical studies, which support that inequality drives innovation and not the opposite. For instance in Table 6 many measures of innovation are not significant. To provide evidence strengthening our claim we have included IV regressions in our analysis. In Table 11 we present our results after we add as an instrument the charges of intellectual property rights. We see that now all our measures of innovation are significant and have a positive effect on the top 1% income share. For instance a 1% increase in the citations on a five year window is associated with a 0.252% rise of the top income share. Like Aghion et al. (2018) the magnitude of the coefficient of innovation in column 4 is much bigger than the corresponding coefficient in column 4 in Table 6. Aghion et al. (2018) state that a good reason could be the interaction between innovation and competition. Also, a 1% increase in the applications from residents is responsible for a 0.21% rise of the top income share while an equivalent increase in the applications from non residents is associated with a 0.0448% increase of the top income share. The results demonstrate that domestic innovation spurs the inequality more than the foreign innovation. The government size and unemployment rate are also significant and with the expected signs. Both variables are in percentages (between 0-1) and we indicate that a 1% increase in the government size decreases the top 1% income share by 2.148% while the same increase in unemployment rate increases top 1% income share by 0.870% (column 4). The rest of the variables are not significant in most tables.

In contrast with Aghion et al. (2018) we find strong evidence in our sample that government size and unemployment rate have the strongest effects and not the financial sector. This is not surprising if we consider that in our sample 13 out of 32 are European countries. Even though Nickell (1997) states that there are big differences among European countries, we cannot skip the fact that unemployment rate is very high in Europe (Fanti and Gori, 2011) and many European countries (high GDP countries) have passed the optimal level of government size compatible with GDP growth rate maximization (Forte and Magazzino, 2011).

We present IV regressions on different income shares in Table 12. The effect of innovation is significant at least at the 5% level and the magnitudes of the coefficients are bigger in very top income shares. Table 13 regresses the different measures of inequality on our measure of innovation (citation on a 5-year window). Innovation influences positively the top 10% income share but in column 3 there is no effect on the average income share. So, after we subtract the top 1% income share from the top 10% the results are not significant. The results are due to the fact that the top 1% income share is included in the top 10%. Also, we see in column 4 that citations have a negative and significant effect on the Theil index. However, we do not have a significant outcome for the Gini index in column 5. Theil index covers only the industrial pay inequality. Moreover, the magnitude of the coefficient is bigger than on the top 1% income share.

There is the probability the effect of innovation on overall income inequality is negative. In Table 14 we present different lags of innovation. We see that in contrast with the OLS results the innovation effect remains significant only for three years. These findings are similar with the corresponding findings of Aghion et al. (2018).

Finally in Table 15, we use data from our second dataset with the total number of patents (granted and not granted). They are the same quality measures as they were in Table 11. The top 1% of most cited patents and the family size are still significant at 5% but the citations on a 5- and 7-year window are now significant at 10%. The top 1% of most cited patents and the family size cannot take into account the defensive patents. But the citations on a 5- or 7-year window "recognize" the defensive patents because the defensive patents receive less citations than the original patents (Aghion at el., 2018).

In the Appendix first we present the results from IV regressions with the second quality measure of innovation, i.e. family size. Next, we provide tables from our second dataset with the total number of patents as robustness check. Specifically, Table 16 presents the results of family size on different measures of inequality. We see that the effect of innovation is positive and significant again on top 1% and top 10% income share but not on the Average Top. Also like citations, there is a negative correlation between family size and the Theil index. In Table 17, we regress family size on different top income shares. The magnitude of the coefficient is again bigger on the top 0.5% and top 0.1% income share.

Next, we use our second sample which includes also defensive patents. In Table 18 we use innovation on different measures of inequality and in Table 19 on different income shares. We can see that the effect of innovation on top income shares is less significant when we apply the full sample of patents.

6 Conclusions

To the best of our knowledge, we make the first attempt to explore the effect of innovation on top income shares at the country level. Also, we have tested the effect of quality measures of innovation with defensive patents on top income shares.

We find strong evidence that innovation boosts top income inequality. We have also checked our findings with different top income shares. Our analysis is based on various quantity[3] and quality[4] measures of innovation. Quality measures of innovation take into consideration the magnitude of the novel inventions in contrast with the quantity measures. As a result, quality measures of innovation have a stronger effect on income inequality. According to our analysis innovation influences inequality for at least 3 years. We have showed also that innovation drives inequality and not the opposite by applying IV analysis. When we tested the effect of innovation on different measures of inequality we found weak evidence that innovation correlates negatively with overall income inequality. We could not explore more the relationship between innovation and top income shares due to the limited data on income shares at the country level. Finally, we have shown that the effect of innovation, when we include defensive patents, is less significant. A future extension of the analysis could include property rights Tebaldi and Elmslie (2013) or taxation Akcigit et al. (2016) as additional control variables.

[3] We use quantity measures like the number of patents granted, and the applications from residents and non residents.

[4] We use quality measures like the number of citations and the family size which each patent belong.

References

Abrams D., Akcigit U., Popadak J., (2013), "Patent Value and Citations: Creative Destruction or Strategic Disruption?" NBER Working Paper No. 19647.

Acemoglu D., Robinson J., Verdier T., (2012), "Can't we all be more like Scandinavians? Asymmetric growth and institutions in an interdependent world", NBER Working Paper No. 18441.

Aghion P., Howitt P., (1992), "A Model of Growth Through Creative Destruction", Econometrica, Vol.60, No.2, 323-351.

Aghion P., Akcigit U., Bergeaud A., Blundell R., Hemous D., (2018), "Innovation and Top Income Inequality", NBER Working Paper No. 21247.

Akcigit U., Baslandze S., Stantcheva S., (2016), "Taxation and the International Mobility of Inventors", American Economic Review, 106(10): 2930-2981.

Alvaredo F., Atkinson A., Chancel L., Piketty T., Saez E., Zucman G. (2017), "Distributional National Accounts (DINA) Guidelines: Concepts and Methods used in WID.world", WID. world working paper series No. 2016/1.

Antonelli C., Gehringer A., (2013), "Innovation and income inequality", Department of Economics and Statistics Cognetti de Martiis. Working Papers 201324, University of Turin.

Barro R. (2008), "Inequality and Growth Revisited",Working Paper Series on Regional Economic Integration no.11, Asian Development Bank.

Bell A., Chetty R., Jaravel X., Petkova N., Reenen J., (2016), "The Lifecycle of Inventors", Working Paper, SSRN Electronic Journal.

Branstetter L., Fisman R., Foley C., (2005), "Do Stronger Intellectual Property Rights Increase International Technology Transfer? Empirical Evidence from U.S. Firm-Level Data", NBER Working Paper No. 11516.

Breau S., Kogler D., Bolton K., (2014), "On the Relationship between Innovation and Wage Inequality: New Evidence from Canadian Cities", Economic Geography 90(4):351–373.

Chu A., Shen G., Zhang X., (2017), "Imports and Intellectual Property Rights on Innovation in China", Fudan University, MPRA Paper No. 81706.

Depalo D., Addario S., (2014), "Shedding Light on Inventors' Returns to Patents" Development Working Papers 375, Centro Studi Luca d'Agliano, University of Milano.

Dutta A., Sharma S., (2008), “Intellectual Property Rights and Innovation in Developing Countries: Evidence from India”, Georgetown University, International Finance Corporation (The World Bank Group).

Fanti L., Gori L., (2011), “A Note on Trade Unions, Unemployment Insurance and Endogenous Growth”, Eastern Economic Journal, 37, (270-280).

Forte F., Magazzino C., (2011), “Optimal Size Government and Economic Growth in EU Countries”, Economica Politica, a. XXVIII, n. 3.

Galbraith J., Halbach B., Shams A., (2013), “Updated Estimates for the World Economy from the University of Texas Inequality Project”, University of Texas, UTIP-UNIDO Data Set.

Hall B. H., Jaffe A. B., Trajtenberg M., (2001), "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." NBER Working Paper 8498.

Hatipoglu O., (2012), “The relationship between inequality and innovative activity: a Schumpeterian theory and evidence from cross-country data”, Scottish Journal of Political Economy, Vol. 59, No. 2.

International Monetary Fund (2017), Balance of Payments Statistics Yearbook and data files.

Kanwar S., Evensont R., (2003), “Does intellectual property protection spur technological change?”, Oxford Economic Papers 55, 235-264.

Lee N., (2011), “Are innovative regions more unequal? Evidence from Europe”, Environment and Planning C: Government and Policy 2011, volume 29, pages 2 – 23.

Lee N., Pose A., (2012), “Innovation and spatial inequality in Europe and USA”, Journal of Economic Geography 13 (2013) pp. 1–22.

Luo W., Pickering A., Monteiro P., (2017), “Inequality and the Size of Government”, Discussion Papers in Economics No. 17/02, University of York.

Nickell S., (1997), “Unemployment and Labor Market Rigidities: Europe versus North America”, Journal of Economic Perspectives, Volume 11, Number 3, Pages 55–7.

OECD Patent Quality Indicators database, March 2018.

OECD REGPAT, database, March 2018.

OECD Triadic Patent Families database, March 2018.

Solt F. (2016), “The Standardized World Income Inequality Database”, Social Science Quarterly 97.

Squicciarini, M., H. Dernis, C. Criscuolo (2013), "Measuring Patent Quality: Indicators of Technological and Economic Value", OECD Science, Technology and Industry Working Papers, 2013/03, OECD Publishing.

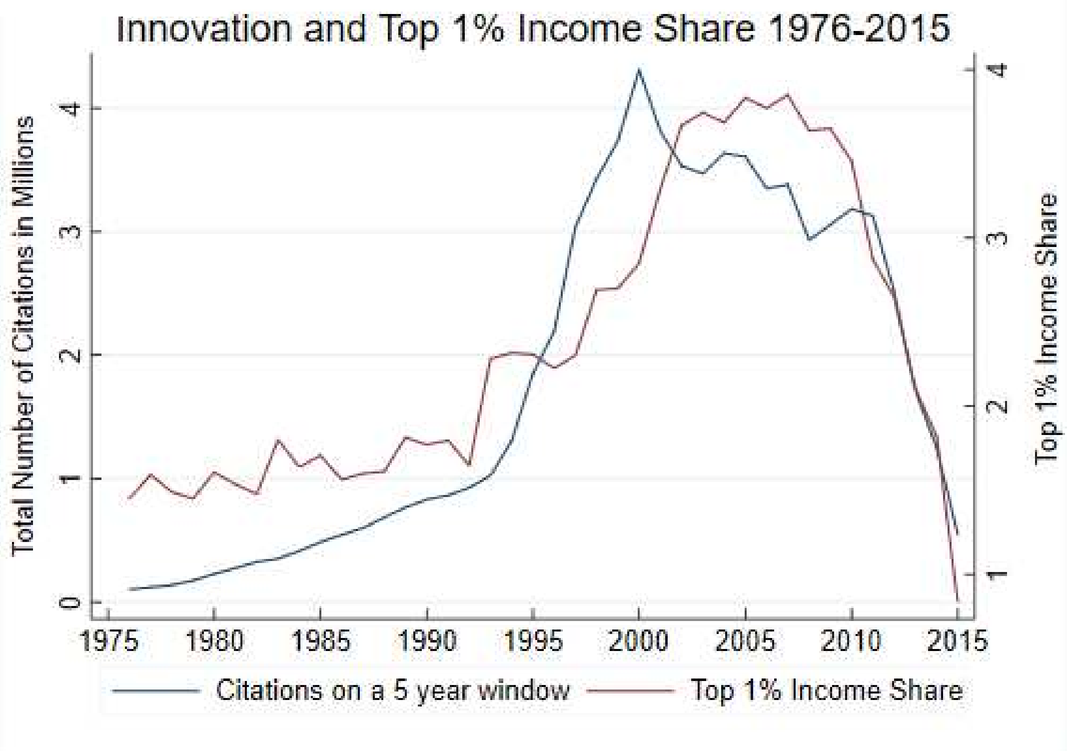
Szymborska H., (2016), "Financial Sector Transformation and Income Inequality- An Empirical Analysis", Financial Internet Quarterly,, e-Finanse" 2016, vol.12/ nr 2, s. 36-48.

Tebaldi E., Elmslie B., (2013), "Does Institutional Quality Impact Innovation? Evidence from Cross-Country Patent Grant Data", Applied Economics, vol. 45, issue 7, 887-900.

The World Bank, World Development Indicators (2018), Data files: World Bank national accounts data.

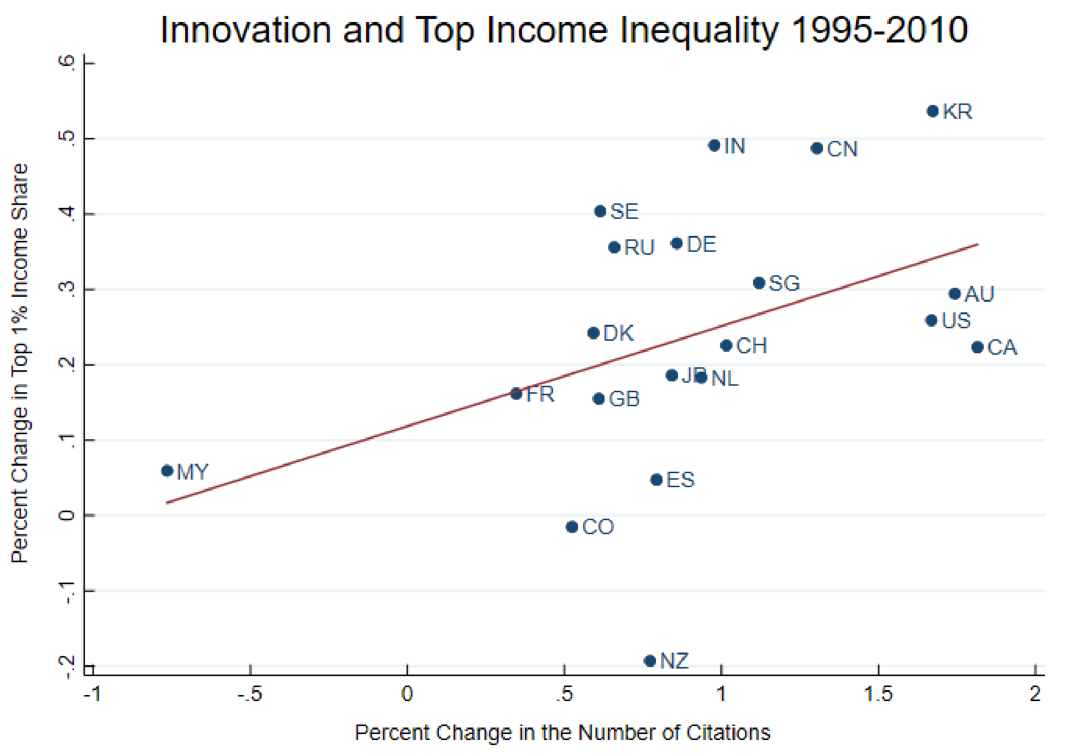
WIPO (2017), "World Intellectual Property Indicators" Geneva, World Intellectual Property Organization.

Figure 1



Notes: Figure 1 plots the number of citations in millions distributed by their year of application against the top 1% income share for the countries as a whole. Observations span the years 1976-2015. Top 1% income shares come from WID and citations data come from the USPTO and EPO.

Figure 2



Notes: Figure 2 plots percentage change in the number of citations per capita against percent change in top 1% income share between 1995 and 2010 for 20 countries. Observations are computed at the country level.

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 10, 1, 0.5 and 0.1) of the income distribution. Time: 1960-2016. Source: WID.
Avgtop	Average income share for the percentiles 10 to 2 in the income distribution. Time: 1960-2016. Source: WID.
Gini	Gini index of inequality with the smallest standard deviation. Time: 1960-2016. Source: SWIID.
Theil	Theil index of inequality. Time: 1963-2015. Source: University of Texas.
	Measures of innovation
Patent	Total patent grants (direct and PCT national phase entries) by filing office. Time: 1980-2015. Source: WIPO.
Applic (N)	Nonresident filings through the Patent Cooperation Treaty procedure or with a national patent office. Time: 1960-2015. Source: WIPO.
Applic (R)	Resident filings through the Patent Cooperation Treaty procedure or with a national patent office. Time: 1960-2015. Source: WIPO.
Cit5	Total number of citations received no longer than 5 years after applications per capita. Time: 1976-2016. Source: OECD.
Cit7	Total number of citations received no longer than 7 years after applications per capita. Time: 1976-2016. Source: OECD.
Famsize	The number of patent offices at which a given invention has been protected per capita. Time: 1976-2016. Source: OECD.
Top1	Number of patents in the top 1% most cited per capita. Time: 1976-2016. Source: OECD.
	Control Variables
Popgr	Growth of total population. Time: 1960-2015. Source: World Bank.
Gvtsize	General government final consumption expenditure (% of GDP). Time: 1960-2016. Source: World Bank.
Unemployment	Unemployment, total (% of total labor force) (national estimate). Time: 1960-2016. Source: World Bank.
Gdppc	Real GDP per capita in US $ (in log). Time: 1960-2016. Source: World Bank.
Finance	Domestic credit provided by financial sector (% of GDP). Time: 1960-2016. Source: World Bank.
	Instrument
Charges	Charges for the use of intellectual property, receipts (BoP, current US$). Time: 1960-2016. Source: International Monetary Fund.

Table 2: Countries sorted by top income share and number of patents granted

Top 1% Income Share			**Patents**
Country		**Country**	
Brazil	0.2769	United States	139206.6
Lebanon	0.2303	Japan	125539.4
Turkey	0.2062	China	63581.03
Zimbabwe	0.1965	Korea	43036.69
Colombia	0.1957	Russian Federation	24429.17
Russian Federation	0.1480	Canada	16460.89
United States	0.1451	Germany	16210.78
Argentina	0.1439	France	14566.72
South Africa	0.1359	Australia	11977.39
India	0.1327	United Kingdom	11050.67
Germany	0.1175	Italy	8546.971
Singapore	0.1164	South Africa	4825.5
Canada	0.1067	Singapore	4625.12
China	0.1052	Spain	3785.139
Indonesia	0.1051	Brazil	3517.875
Malaysia	0.1000	India	3507.6
France	0.0992	New Zealand	3484.389
United Kingdom	0.0980	Sweden	3074.333
Switzerland	0.0967	Switzerland	2957.857
Japan	0.0955	Netherlands	2587.639
Spain	0.0907	Finland	1886.778
Korea	0.0890	Malaysia	1865.774
Ireland	0.0817	Argentina	1708.5
Italy	0.0797	Indonesia	1105
Portugal	0.0762	Denmark	1034.444
New Zealand	0.0745	Ireland	882.4
Finland	0.0725	Turkey	809.0882
Mauritius	0.0712	Portugal	789.5278
Sweden	0.0665	Colombia	545
Netherlands	0.0659	Lebanon	308.4
Australia	0.0652	Zimbabwe	178.7647
Denmark	0.0622	Mauritius	5.6

Notes: The right column illustrates the countries sorted by the mean of top 1 percent income share over the period 1960-2015 while the left column represents the countries sorted by the mean number of patents granted over the period 1980-2015.

Table 3: Summary statistics of the main variables

Variable	Mean	Std. Dev.	Min	Max
Measures of Inequality				
Top 10% Income Share	13109.01	27272.39	0.1396	0.6508
Avgtop	22734.3	69637.22	0.0100	0.0510
Top 1% Income Share	17705.88	42811.72	0.0345	0.2943
Top 0.5% Income Share	59959.04	241493.1	0.0198	0.2411
Top 0.1% Income Share	79754.72	323702.8	0.0054	0.1661
Theil	59816.01	159970.8	0.0013	0.1483
Gini	78.63667	291.099	0.2971	0.6954
Measures of Innovation				
Patents	0.0362	0.0254	1	359316
Applic(N)	0.0760	0.0397	3	301075
Applic(R)	0.1046	0.0478	1	968252
Cit5	0.3380	0.0750	0	2581189
Cit7	0.0265	0.0044	0	3449904
Famsize	0.0313	0.0244	1	1183832
Top 1% of Citations	0.4603	0.0632	0	2381
Rest of the control variables				
Popgrowth	1.1418	0.9208	-1.8537	7.0610
Gdppc(log)	9.4405	1.2725	4.8825	11.2359
Unemployment	6.9325	4.3912	0.08	27.14
Finance	89.0233	60.3729	7.1173	363.306
Gonverment	15.9056	4.7387	2.9755	27.935
Instrument				
Charges	3.63e+09	1.29e+10	0	1.30e+11

Notes: Summary statistics for the main variables calculated over the period 1960-2016. GDP per capita is calculated in $ per capita.

Table 4: Descriptive statistics of measures of inequality

1990 Measures of Inequality	Mean	Min	Max	P5	P25	P50	P90
Top 1% Income Share	0.0843	0.0493	0.1454	0.0517	0.0664	0.0810	0.1122
Top 10% Income Share	0.3052	0.1555	0.3892	0.1957	0.2687	0.3108	0.3781
Avgtop	0.0248	0.0118	0.0301	0.0118	0.0226	0.0255	0.0299
Top 0.5% Income Share	0.0585	0.0331	0.1095	0.0337	0.04	0.0575	0.0877
Top 0.1% Income Share	0.0269	0.0109	0.0555	0.0109	0.016	0.0269	0.0545
Theil	0.0296	0.0029	0.0719	0.0062	0.0127	0.0258	0.0562
Gini	0.4502	0.3092	0.6330	0.3814	0.4163	0.4428	0.5061
2010 Measures of Inequality	Mean	Min	Max	P5	P25	P50	P90
Top 1% Income Share	0.1382	0.0641	0.2776	0.0645	0.0898	0.1255	0.2122
Top 10% Income Share	0.3953	0.2463	0.6067	0.2688	0.3099	0.3966	0.5217
Avgtop	0.0290	0.0169	0.0468	0.0227	0.0245	0.0290	0.0347
Top 0.5% Income Share	0.1029	0.0403	0.2270	0.0432	0.0643	0.0923	0.1628
Top 0.1% Income Share	0.0577	0.0186	0.1347	0.0220	0.0322	0.0484	0.1019
Theil	0.0344	0.0069	0.0919	0.0132	0.0153	0.0265	0.0829
Gini	0.4862	0.3446	0.6954	0.3993	0.4531	0.4822	0.5441

Notes: Summary statistics includes mean, percentile thresholds, minimum and maximum for our seven measures of inequality.

Table 5: Descriptive statistics of measures of innovation

1990 Measure of Innovation	Mean	Min	Max	P5	P25	P50	P90
Patents	10432.81	13	90366	134	852	3402.5	19073
Applic(N)	8963.88	3	80520	764	1843	4001	24375
Applic(R)	21758.76	1	332952	92	955	2218	30724
Cit5	26777.35	3	421495	5	42	556	25119
Cit7	40939.26	3	668192	7	72	805	37552
Famsize	51963.65	6	583991	21	187	1547	110387
Top 1% of Citations	73.1613	0	1360	0	0	1	52
2010 Measure of Innovation	Mean	Min	Max	P5	P25	P50	P90
Patents	28304.29	8	222693	140	1144.5	4394.5	135110
Applic(N)	20776.53	46	248249	59	353	5137	46406.5
Applic(R)	37691.6	133	293066	499	1231	2853.5	186891
Cit5	102712.1	44	1657790	63	610	9587	110670
Cit7	108812.5	44	1763210	74	638	9957	118880
Famsize	85133.87	35	875996	84	1580	17523	163576
Top 1% of Citations	104.0968	0	1915	0	0	4	79

Notes: Summary statistics includes mean, percentile thresholds, minimum and maximum for our seven measures of innovation.

Table 6: Top 1% income share and innovation

Dependent Variable	Top 1% income share						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Measure of Innovation	Patents	Applic (N)	Applic (R)	Cit5	Cit7	Famsize	Top1
Innovation	0.0117	0.00777	-0.100***	0.0315**	0.0263*	0.0240	0.0364***
	(1.57)	(1.21)	(-8.96)	(2.55)	(1.93)	(1.27)	(5.00)
Popgr	0.0285**	0.00853	-0.00648	0.0168	0.0176	0.0174	0.0176
	(2.05)	(0.70)	(-0.62)	(1.45)	(1.51)	(1.48)	(1.64)
Gvtsize	-2.734***	-3.247***	-3.562***	-2.802***	-2.817***	-2.846***	-2.665***
	(-5.77)	(-7.17)	(-8.76)	(-6.47)	(-6.51)	(-6.66)	(-6.27)
Unemployment	0.918***	0.577**	0.940***	0.965***	0.976***	0.971***	1.086***
	(3.45)	(2.09)	(3.78)	(3.92)	(3.94)	(3.92)	(4.60)
Gdppc	0.0954***	0.0331	0.365***	0.0901***	0.0918***	0.0925***	0.184***
	(2.80)	(1.24)	(9.25)	(4.03)	(4.12)	(4.22)	(7.19)
Finance	0.0204	0.0293	0.0156	0.0410*	0.0413*	0.0425*	0.0260
	(0.80)	(1.14)	(0.67)	(1.69)	(1.69)	(1.72)	(1.12)
R^2	0.938	0.919	0.927	0.932	0.932	0.931	0.935
Observations	652	777	774	737	737	737	737

Notes: Innovation is taken in logs and lagged by one year. The dependent variable is the log of the top 1% income share. Panel data OLS regressions with country and year fixed effects. Time span for innovation: 1982-2015 for column 1, 1962-2015 for columns 2 and 3 and 1978-2015 for columns 4-7. Autocorrelation and heteroskedasticity robust standard errors are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 7: Top 1% income share and innovation

Dependent Variable	Top 1% income share						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Measure of Innovation	Patents	Applic (N)	Applic (R)	Cit5	Cit7	Famsize	Top1
Innovation	0.0117	0.00777	-0.100***	0.0315*	0.0263	0.0240	0.0364***
	(0.73)	(0.35)	(-4.56)	(1.96)	(1.54)	(1.10)	(3.19)
Popgr	0.0285	0.00853	-0.00648	0.0168	0.0176	0.0174	0.0176
	(1.13)	(0.30)	(-0.44)	(0.60)	(0.63)	(0.60)	(0.65)
Gvtsize	-2.734***	-3.247***	-3.562***	-2.802***	-2.817***	-2.846***	-2.665***
	(-3.10)	(-3.18)	(-4.19)	(-3.23)	(-3.22)	(-3.20)	(-3.00)
Unemployment	0.918	0.577	0.940*	0.965	0.976	0.971	1.086*
	(1.40)	(0.96)	(1.88)	(1.66)	(1.67)	(1.63)	(2.00)
Gdppc	0.0954	0.0331	0.365***	0.0901	0.0918	0.0925	0.184***
	(1.51)	(0.45)	(4.59)	(1.41)	(1.44)	(1.50)	(4.13)
Finance	0.0204	0.0293	0.0156	0.0410	0.0413	0.0425	0.0260
	(0.36)	(0.54)	(0.35)	(0.71)	(0.72)	(0.72)	(0.51)
R^2	0.938	0.919	0.927	0.932	0.932	0.931	0.935
Observations	652	777	774	737	737	737	737

Notes: Innovation is taken in logs and lagged by one year. The dependent variable is the log of the top 1% income share. Panel data OLS regressions with country and year fixed effects. Time span for innovation: 1982-2015 for column 1, 1962-2015 for columns 2 and 3 and 1978-2015 for columns 4-7. Autocorrelation and heteroskedasticity robust standard errors clustered at the country level 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	Top10%	Top1%	Top0.5%	Top0.1%
	(1)	**(2)**	**(3)**	**(4)**
Measure of Innovation	Cit5	Cit5	Cit5	Cit5
Innovation	0.0113	0.0248*	0.0308*	0.0461**
	(1.39)	(1.77)	(1.89)	(2.07)
Popgr	-0.0206***	-0.0111	-0.00707	0.00701
	(-3.11)	(-1.07)	(-0.57)	(0.43)
Gvtsize	-0.184	-2.370***	-2.967***	-4.315***
	(-0.79)	(-5.35)	(-5.54)	(-5.53)
Unemployment	0.642***	0.695**	0.621*	0.690
	(4.27)	(2.48)	(1.85)	(1.41)
Gdppc	0.110***	0.0888***	0.0766***	0.0640**
	(10.83)	(3.99)	(3.03)	(2.07)
Finance	-0.0191	-0.0299	-0.0347	-0.0345
	(-1.43)	(-1.22)	(-1.19)	(-0.84)
R^2	0.950	0.947	0.942	0.930
Observations	588	588	588	588

Notes: Innovation is taken in logs and lagged by one year. The dependent variables are taken in logs. Panel data OLS regressions with country and year fixed effects. Time span for innovation: 1978-2015. Autocorrelation and heteroskedasticity robust standard errors are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 9: Innovation and various measures of inequality

Dependent Variable	Top1%	Top10%	Avgtop	Theil	Gini
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Measure of Innovation	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation	0.0315**	0.0103	0.00687	-0.00145	0.00136
	(2.55)	(1.07)	(0.76)	(-0.06)	(0.32)
Popgr	0.0168	-0.00751	-0.0105	-0.0561*	0.00535
	(1.45)	(-1.13)	(-1.51)	(-1.92)	(1.55)
Gvtsize	-2.802***	-0.364	0.833***	0.0127	-0.196*
	(-6.47)	(-1.43)	(3.25)	(0.02)	(-1.67)
Unemployment	0.965***	0.766***	0.785***	1.045**	0.517***
	(3.92)	(5.43)	(5.77)	(2.11)	(6.78)
Gdppc	0.0901***	0.112***	0.114***	-0.0589	0.0522***
	(4.03)	(10.95)	(10.34)	(-0.39)	(2.67)
Finance	0.0410*	0.0251*	0.0292*	0.208***	0.0358***
	(1.69)	(1.72)	(1.94)	(4.66)	(5.84)
R^2	0.932	0.932	0.889	0.873	0.867
Observations	737	680	676	868	962

Notes: Innovation is taken in logs and lagged by one year. The dependent variables are taken in logs. Panel data OLS regressions with country and year fixed effects. Time span for innovation: 1978-2015 for columns 1- 4 and 1978-2016 for column 5. Autocorrelation and heteroskedasticity robust standard errors are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1levels of significance.

Table 10: Top 1% income share and innovation at different lags

Dependent Variable	Top 1% Income Share					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Cit5	Cit5	Cit5	Cit5	Cit5	Cit5
Lag of Innovation	1 year	2 years	3 years	4 years	5 years	6 years
Innovation	0.0323**	0.0466***	0.0384***	0.0344***	0.0324**	0.0351***
	(2.16)	(3.03)	(3.07)	(2.81)	(1.99)	(2.78)
Popgr	0.0115	0.0123	0.0103	0.00989	0.0109	0.0132
	(0.92)	(0.94)	(0.80)	(0.79)	(0.88)	(1.05)
Gvtsize	-2.345***	-2.201***	-2.244***	-2.241***	-2.275***	-2.353***
	(-4.56)	(-4.36)	(-4.39)	(-4.38)	(-4.36)	(-4.71)
Unemployment	0.896***	0.856***	0.852***	0.850***	0.832***	0.843***
	(3.39)	(3.21)	(3.23)	(3.23)	(3.26)	(3.23)
Gdppc	0.159***	0.151***	0.151***	0.149***	0.153***	0.147***
	(5.88)	(5.29)	(5.57)	(5.36)	(5.69)	(5.39)
Finance	0.0159	0.0101	0.0137	0.0159	0.0144	0.0124
	(0.65)	(0.41)	(0.55)	(0.64)	(0.58)	(0.50)
R^2	0.938	0.939	0.939	0.939	0.939	0.939
Observations	652	652	652	652	652	652

Notes: Innovation is taken in logs. The lag between the dependent variable and the innovation measures ranges from 1 year to 6 years. The dependent variable is taken in log. Panel data OLS regressions with country and year fixed effects. Time span for innovation: 1983-2015. Autocorrelation and heteroskedasticity robust standard errors are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 11: Regression of innovation on top 1% income share – IV estimation

Dependent Variable	Top 1% Income Share						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Measure of Innovation	Patents (G)	Applic (N)	Applic (R)	Cit5	Cit7	Famsize	Top1
Innovation	0.0331*	0.0448***	0.210***	0.252**	0.256**	0.181**	0.129**
	(1.86)	(4.32)	(2.81)	(1.98)	(1.99)	(2.39)	(2.47)
Popgr	0.00821	0.00319	0.00586	-0.00578	-0.000514	-0.00772	-0.00899
	(0.66)	(0.24)	(0.21)	(-0.46)	(-0.04)	(-0.66)	(-0.81)
Gvtsize	-2.108***	-1.825***	-1.932***	-2.148***	-2.216***	-2.248***	-1.688***
	(-4.14)	(-3.15)	(-2.59)	(-3.55)	(-3.59)	(-4.59)	(-2.81)
Unemployment	0.928***	0.745**	0.654	0.870**	0.947***	1.066***	1.145***
	(2.84)	(2.36)	(1.52)	(2.46)	(2.74)	(3.45)	(4.18)
Gdppc	0.000339	-0.0782**	-0.627***	-0.0122	-0.0131	0.0226	0.381***
	(0.01)	(-2.13)	(-2.62)	(-0.21)	(-0.22)	(0.62)	(3.02)
Finance	-0.00820	-0.00518	0.0176	-0.0166	-0.0197	0.00794	-0.0551
	(-0.27)	(-0.16)	(0.36)	(-0.40)	(-0.47)	(0.24)	(-1.14)
F first stage	117.03	243.82	40.49	10.90	11.04	48.48	13.38
R^2	0.653	0.635	0.332	0.513	0.501	0.656	0.609
Observations	554	617	614	616	616	616	616

Notes: Innovation is taken in logs and lagged by one year. Panel data IV 2SLS regressions with country and year fixed effects. Innovation is instrumented by the charges for the use of intellectual property. The lead between the instrument and the endogenous variable is set to 1 year. Time span for innovation: 1981-2014 for column 1, 1969-2014 for columns 2, 3 and 1977-2014 for columns 4-7. Number of groups: 30 for column 1 and 31 for columns 2-7. Autocorrelation and heteroskedasticity robust standard errors are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 12: Regression of innovation on different income shares – IV estimation

Dependent Variable	Top10%	Top1%	Top0.5%	Top0.1%
	(1)	**(2)**	**(3)**	**(4)**
Measure of Innovation	Cit5	Cit5	Cit5	Cit5
Innovation	0.209***	0.252**	0.319**	0.274**
	(2.71)	(1.98)	(2.33)	(2.02)
Popgr	-0.00595	-0.00578	-0.00844	-0.0113
	(-0.60)	(-0.46)	(-0.62)	(-0.73)
Gvtsize	0.150	-2.148***	-2.161***	-4.092***
	(0.35)	(-3.55)	(-2.75)	(-4.16)
Unemployment	0.721***	0.870**	0.964**	1.253**
	(3.06)	(2.46)	(2.28)	(2.31)
Gdppc	0.0404	-0.0122	-0.0267	-0.0222
	(1.14)	(-0.21)	(-0.40)	(-0.37)
Finance	-0.00419	-0.0166	-0.0878*	-0.103*
	(-0.17)	(-0.40)	(-1.83)	(-1.73)
F first stage	15.95	10.90	12.43	14.41
R^2	0.163	0.513	0.423	0.613
Observations	565	616	579	539

Notes: Innovation is taken in logs and lagged by one year. Panel data IV 2SLS regressions with country and year fixed effects. Innovation is instrumented by the charges for the use of intellectual property. The lead between the instrument and the endogenous variable is set to 1 year. The dependent variables are taken in logs. Time span for innovation: 1977-2014 for column 1-4. Number of groups: 27 for columns 1 and 4, 31 for column 2 and 29 for column 3. Autocorrelation and heteroskedasticity robust standard errors are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 13: Regression of innovation on different measures of inequality – IV estimation

Dependent Variable	Top1%	Top10%	Avgtop	Theil	Gini
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Measure of Innovation	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation	0.252**	0.209***	0.0592	-0.383*	0.0273
	(1.98)	(2.71)	(1.27)	(-1.66)	(0.73)
Popgr	-0.00578	-0.00595	-0.00384	0.0392	0.00689*
	(-0.46)	(-0.60)	(-0.47)	(1.13)	(1.85)
Gvtsize	-2.148***	0.150	0.756**	-1.324	-0.290**
	(-3.55)	(0.35)	(2.55)	(-1.40)	(-2.31)
Unemployment	0.870**	0.721***	0.808***	1.832**	0.498***
	(2.46)	(3.06)	(4.64)	(2.22)	(5.59)
Gdppc	-0.0122	0.0404	0.100***	0.479**	0.0532**
	(-0.21)	(1.14)	(5.34)	(2.44)	(2.23)
Finance	-0.0166	-0.00419	0.0362*	0.367***	0.0369***
	(-0.40)	(-0.17)	(1.77)	(4.56)	(4.61)
F first stage	10.90	15.95	17.46	10.90	6.73
R^2	0.513	0.163	0.361	0.023	0.426
Observations	616	565	561	734	828

Notes: Innovation is taken in logs and lagged by one year. Panel data IV 2SLS regressions with country and year fixed effects. Innovation is instrumented by the charges for the use of intellectual property. The lead between the instrument and the endogenous variable is set to 1 year. The dependent variables are taken in logs. Time span for innovation: 1977-2014 for column 1-5. Number of groups: 31 for columns 1, 4, 5 and 27 for columns 2 and 3. Autocorrelation and heteroskedasticity robust standard errors are presented 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 at different lags – IV estimation

Dependent Variable	Top 1% Income Share				
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Measure of Innovation	Cit5	Cit5	Cit5	Cit5	Cit5
Lag of Innovation	1 year	2 years	3 years	4 years	5 years
Innovation	0.252**	0.204*	0.230*	0.270	0.386
	(1.98)	(1.94)	(1.76)	(1.62)	(0.99)
Popgr	-0.00578	-0.00590	-0.0170	-0.0297	-0.0201
	(-0.46)	(-0.42)	(-1.16)	(-1.49)	(-1.28)
Gvtsize	-2.148***	-1.439**	-0.718	-0.399	-0.128
	(-3.55)	(-2.22)	(-0.79)	(-0.41)	(-0.07)
Unemployment	0.870**	0.611	0.481	0.519	0.231
	(2.46)	(1.40)	(1.00)	(1.06)	(0.24)
Gdppc	-0.0122	0.00945	0.00223	-0.0235	-0.0423
	(-0.21)	(0.19)	(0.04)	(-0.27)	(-0.25)
Finance	-0.0166	-0.0204	-0.0391	-0.0478	-0.0932
	(-0.40)	(-0.50)	(-0.84)	(-1.00)	(-0.96)
F first stage	10.90	11.05	8.77	5.32	1.34
R^2	0.513	0.579	0.550	0.450	0.226
Observations	616	603	596	585	571

Notes: Innovation is taken in logs. The lag between the dependent variable and the innovation measure ranges from 1 year to 5 years. Panel data IV 2SLS regressions with country and year fixed effects. Innovation is instrumented by the charges for the use of intellectual property. The lead between the instrument and the endogenous variable is set to 1 year. The dependent variable is taken in log. Time span for innovation: 1977-2014 for column 1, 1978-2014 for column 2, 1979-2014 for column 3, 1980-2014 for column 4 and 1981-2014 for column 5. Number of groups: 31 for columns 1, 2, 3, 4 and 5. Autocorrelation and heteroskedasticity robust standard errors are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 15: Regression of innovation on top 1% income share – IV estimation

Dependent Variable	Top 1% Income Share			
	(1)	**(2)**	**(3)**	**(4)**
Measure of Innovation	Cit5	Cit7	Famsize	Top1
Innovation	0.287*	0.291*	0.176**	0.133**
	(1.82)	(1.83)	(2.48)	(2.43)
Popgr	-0.00266	0.00392	-0.00448	-0.00718
	(-0.20)	(0.30)	(-0.39)	(-0.62)
Gvtsize	-2.059***	-2.160***	-2.323***	-1.574**
	(-3.16)	(-3.27)	(-4.79)	(-2.49)
Unemployment	0.788**	0.894**	1.081***	1.093***
	(2.07)	(2.49)	(3.58)	(3.87)
Gdppc	0.00840	0.00875	0.0306	0.399***
	(0.15)	(0.16)	(0.91)	(2.93)
Finance	-0.0226	-0.0258	0.00486	-0.0761
	(-0.50)	(-0.56)	(0.15)	(-1.35)
F first stage	7.61	7.64	70.10	13.66
R^2	0.476	0.461	0.667	0.607
Observations	614	614	614	614

Notes: Innovation is taken in logs and lagged by one year. Panel data IV 2SLS regressions with country and year fixed effects. Innovation is instrumented by the charges for the use of intellectual property. The lead between the instrument and the endogenous variable is set to 1 year. Time span for innovation: 1977-2014 for columns 1-4. Number of groups: 31 for columns 1-4. Autocorrelation and heteroskedasticity robust standard errors are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Appendix

Methodology of construction of the database

Every patent to our database has two unique id codes. The first is an id code for the PATSTAT database. The second is the application number of each patent. These codes are important because with their help we manage to assign the patents to their inventors. We allocate the patents based on the inventors' location (country). According to Aghion et al. (2018) the location of the inventor and assignee coincide (they have a correlation higher than 95%). The locations of the inventors for the European Patent Office exist in the OECD REGPAT Database. The OECD REGPAT Database has also the id code of the PATSTAT database. We match the patents with the inventors and their countries. We manage to assign 1,617,825 out of 3,190,373 patents to their inventors. To extend our sample of patents we use also the locations of the inventors of the United States Patent and Trademark Office. First, we use the id of the PATSTAT Database to match the patents with the inventors who belong in the USPTO and exist in the OECD Database on Triadic Patent Families. We have managed to match 2,081,862 out of 7,763,046 patents from the Patent Quality Indicators Database with their inventors. To avoid losing the rest of the observations we use the Database from the Patent Data Project of National Bureau of Economic Research. This Database has the inventors of USPTO from 1901 to 2006. We use this time the USPTO application number of each patent to do the matching. This method allows us to save 2,203,978 more patents. When we combine these datasets our sample includes 5,903,665 different patents from EPO and USPTO. After we allocate the different patents based on their inventors to the countries, we have a sample of 15,454,398 patents for 32 countries over the period 1976-2015.

Table 16: Regression of innovation on different measures of inequality – IV estimation

Dependent Variable	Top1%	Top10%	Avgtop	Theil	Gini
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Measure of Innovation	Famsize	Famsize	Famsize	Famsize	Famsize
Innovation	0.181**	0.162***	0.0482	-0.359*	0.0249
	(2.39)	(3.32)	(1.31)	(-1.83)	(0.74)
Popgr	-0.00772	-0.0147	-0.00582	0.0228	0.00762**
	(-0.66)	(-1.45)	(-0.71)	(0.66)	(2.12)
Gvtsize	-2.248***	-0.558*	0.599**	0.277	-0.373***
	(-4.59)	(-1.72)	(2.10)	(0.30)	(-2.94)
Unemployment	1.066***	0.708***	0.779***	1.117**	0.511***
	(3.45)	(3.44)	(4.40)	(2.17)	(6.26)
Gdppc	0.0226	0.0580**	0.105***	0.363**	0.0579***
	(0.62)	(2.23)	(6.26)	(2.20)	(2.90)
Finance	0.00794	0.0171	0.0430**	0.280***	0.0405***
	(0.24)	(0.81)	(2.21)	(4.90)	(5.72)
F first stage	48.48	56.81	55.98	27.93	20.46
Observations	616	565	561	734	828
R^2	0.656	0.511	0.390	0.126	0.474

Notes: Innovation is taken in logs and lagged by one year. Panel data IV 2SLS regressions with country and year fixed effects. Innovation is instrumented by the charges for the use of intellectual property. The lead between the instrument and the endogenous variable is set to 1 year. The dependent variables are taken in logs. Time span for innovation: 1977-2014 for column 1-5. Number of groups: 31 for columns 1, 4, 5 and 27 for columns 2 and 3. Autocorrelation and heteroskedasticity robust standard errors are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 17: Regression of innovation on different income shares – IV estimation

Dependent Variable	Top10%	Top1%	Top0.5%	Top0.1%
	(1)	**(2)**	**(3)**	**(4)**
Measure of Innovation	Famsize	Famsize	Famsize	Famsize
Innovation	0.162***	0.181**	0.228***	0.193**
	(3.32)	(2.39)	(2.82)	(2.21)
Popgr	-0.0147	-0.00772	-0.0148	-0.0192
	(-1.45)	(-0.66)	(-1.13)	(-1.21)
Gvtsize	-0.558*	-2.248***	-2.684***	-4.486***
	(-1.72)	(-4.59)	(-5.09)	(-5.82)
Unemployment	0.708***	1.066***	1.039***	1.214**
	(3.44)	(3.45)	(2.98)	(2.41)
Gdppc	0.0580**	0.0226	0.0187	0.0153
	(2.23)	(0.62)	(0.47)	(0.39)
Finance	0.0171	0.00794	-0.0391	-0.0699
	(0.81)	(0.24)	(-1.20)	(-1.46)
F first stage	56.81	48.48	55.05	68.86
R^2	0.511	0.656	0.658	0.683
Observations	565	616	579	539

Notes: Innovation is taken in logs and lagged by one year. Panel data IV 2SLS regressions with country and year fixed effects. Innovation is instrumented by the charges for the use of intellectual property. The lead between the instrument and the endogenous variable is set to 1 year. The dependent variables are taken in logs. Time span for innovation: 1977-2014 for column 1- 4. Number of groups: 27 for columns 1 and 4, 31 for column 2 and 29 for column 3. Autocorrelation and heteroskedasticity robust standard errors are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 18: Regression of innovation on different measures of inequality – IV estimation

Dependent Variable	Top1%	Top10%	Avgtop	Theil	Gini
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Measure of Innovation	Cit5	Cit5	Cit5	Cit5	Cit5
Innovation	0.287*	0.215***	0.0609	-0.486*	0.0405
	(1.82)	(2.64)	(1.27)	(-1.65)	(0.92)
Popgr	-0.00266	-0.00580	-0.00375	0.0347	0.00746**
	(-0.20)	(-0.57)	(-0.45)	(0.95)	(1.98)
Gvtsize	-2.059***	0.244	0.781**	-1.505	-0.274**
	(-3.16)	(0.55)	(2.55)	(-1.46)	(-2.02)
Unemployment	0.788**	0.715***	0.807***	1.915**	0.503***
	(2.07)	(3.07)	(4.62)	(2.18)	(5.76)
Gdppc	0.00840	0.0609**	0.106***	0.473**	0.0526**
	(0.15)	(2.07)	(6.92)	(2.33)	(2.26)
Finance	-0.0226	-0.00474	0.0360*	0.362***	0.0373***
	(-0.50)	(-0.18)	(1.72)	(4.48)	(4.81)
F first stage	7.61	14.54	15.87	7.71	5.21
R^2	0.476	0.190	0.357	-0.030	0.392
Observations	614	563	559	717	809

Notes: Innovation is taken in logs and lagged by one year. Panel data IV 2SLS regressions with country and year fixed effects. Innovation is instrumented by the charges for the use of intellectual property. The lead between the instrument and the endogenous variable is set to 1 year. The dependent variables are taken in logs. Time span for innovation: 1977-2014 for column 1-5. Number of groups: 31 for columns 1, 4, 5 and 27 for columns 2 and 3. Autocorrelation and heteroskedasticity robust standard errors are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.

Table 19: Regression of innovation on different income shares – IV estimation

Dependent Variable	Top10%	Top1%	Top0.5%	Top0.1%
	(1)	**(2)**	**(3)**	**(4)**
Measure of Innovation	Cit5	Cit5	Cit5	Cit5
lcits5_pcl1	0.215***	0.287*	0.362**	0.295*
	(2.64)	(1.82)	(2.15)	(1.94)
Popgr	-0.00580	-0.00266	-0.00479	-0.00900
	(-0.57)	(-0.20)	(-0.33)	(-0.56)
Gvtsize	0.244	-2.059***	-2.024**	-4.126***
	(0.55)	(-3.16)	(-2.36)	(-4.12)
Unemployment	0.715***	0.788**	0.897**	1.251**
	(3.07)	(2.07)	(2.08)	(2.34)
Gdppc	0.0609**	0.00840	0.00118	0.00733
	(2.07)	(0.15)	(0.02)	(0.15)
Finance	-0.00474	-0.0226	-0.0948*	-0.112*
	(-0.18)	(-0.50)	(-1.79)	(-1.76)
F first stage	14.54	7.61	8.98	11.40
N	563	614	577	537
R^2	0.190	0.476	0.367	0.609

Notes: Innovation is taken in logs and lagged by one year. Panel data IV 2SLS regressions with country and year fixed effects. Innovation is instrumented by the charges for the use of intellectual property. The lead between the instrument and the endogenous variable is set to 1 year. The dependent variables are taken in logs. Time span for innovation: 1977-2014 for column 1-4. Number of groups: 27 for columns 1 and 4, 31 for column 2 and 29 for column 3. Autocorrelation and heteroskedasticity robust standard errors are presented in parentheses. ***, ** and * respectively indicate 0.01, 0.05 and 0.1 levels of significance.