Better safe than sorry?

Infrastructure, State Incentives and Innovation across U.S states

JOB MARKET PAPER

Santiago Caram*

This version: March 2022

First version: September 2019

Abstract

In light of the current U.S business slowdown in a context where ideas are difficult to find and "superstar firms" concentrate almost all knowledge and technology, two questions arise: Can State incentives monetary (R&D tax-credits) and non-monetary (the diffusion of past public policies) coupled with infrastructure investments (public and private) boost job creation rates among industries? If so, in which sectors these effects will be larger? Would it be any effect on wages? This paper examines the effectiveness of the above-mentioned policies over U.S labor markets. To this end, I propose a simple regional model in which workers have some kind of mobility and firms make use of both types of incentives. Then, I identify some stylized facts from the U.S business slowdown (e.g., higher knowledge gap and lower competition between public firms) in high-tech sectors. I focus on these group of firms because regional economies are expected to be benefited more from the increasing value share generated as well as different forms of agglomeration economies and thick market externalities. I find that both types of incentives coupled with private non-residential expenditures have a positive business "enabler" entrepreneurship effect on employment rates in top value-added sectors while highway expenditures display a negative "disabler" effect but only restricted to sectors with lower value share. Interestingly, generous State subsidies and R&D user cost are positively associated to higher job creation rates in bottom rather than top value-added sectors. When spatial effects are accounted for, I find positive and significant spillover effects of R&D tax-credits in both groups provided that efficiency gains are pass-through to more patents. Likewise, in the case of wages the same trends from employment are observed. That is the median worker in lower value-added sectors benefits more from these public policies as his/her salary is higher compared to workers in top value-added sectors. Evidently, larger amounts of knowledge in sectors with increasing value shares do not necessarily imply more jobs and/or better salaries. Therefore, the findings of this paper are an important departure from the "beggar the neighbour effect" and the effectiveness of traditional public policies over labor markets.

---

*I am indebted to an American professor for the insights about regional economies and the encouragement. Moreover, I have been benefited from many fruitful suggestions of former professors in Argentina and the U.S as well as stimulating discussions with friends who were also pursuing their Ph.D's. I bear sole responsibility for the content or errors of the paper and of course, the opinions expressed here do not represent the view of the University of Messina. The usual disclaimer applies. Email: scaram@unime.it

"If you look at history, innovation doesn't come just from giving people incentives; it comes from creating environments where their ideas can connect."

Steve Johnson, Science author & media theorist

1 Introduction

Regional development as well as the identification of the proper policies have always captured the attention of scholars. In this vein, infrastructure investments —whether these are performed by public or private hands— may be regarded as "boosters" for regional and local economies in terms of new establishments, additional jobs and foremost the enhancement of distressed places. Although the scope of infrastructure has evolved over time, labor market response at an industry level has not been fully explored.[1]

At first glance, the transmission mechanism by which these expenditures affect regional employment constitute a first-order "direct effect" regardless the industry under consideration. For instance, during the Great Recession in the U.S, high-tech industries like computer and electronic manufacturing experienced a fell of 6.2 percent jobs in Colorado with an overall fall of 3.6 percent jobs in total private employment whilst in the telecommunication sector, the converse occurred: private employment grew 4.6 and 0.7 percent respectively.[2]

In response to such disparities in regional performance, policymakers may consider that well-targeted stimulus spending may facilitate recovery for the areas most in need (e.g., the American Recovery and Reinvestment Act better known as "ARRA").[3] Although, the interaction of several and possibly opposed effects within each region (e.g., different tax incentives, innovation activities performed by firms, quality of institutions, etc) might affect the final outcome; ultimately, whether or not is there any effect over employment remains an empirical question.

Studies show that the high-tech sector has the largest multiplier effect (5.9) in terms of job-creation rates whereas the effect on manufacturing industries is positive but far modest (2.6).

[1] For example, a recent work of Hooper et al. (2018) explores the relationship between transportation spending (i.e., highways) and income inequality across U.S states. Results using top income shares and the Gini index, show that the construction sector plays an important role in reducing inequality. Hence, infrastructure investments within this industry allow workers to switch to better remunerated jobs.

[2] Retrieved from Quarterly Census of Employment and Wages (QCEW) on July 10th 2019.

[3] The empirical results considering the ARRA event study cannot reject the "crowding out" hypothesis (Dupor, 2017). In fact, after controlling for population density, hardly any positive or negative correlation exists among U.S states. More precisely, it seems that highway infrastructure spending may have been very similar with or without the implementation of such program.

Expressed differently, for each additional job in the former industry, 5.9 jobs are created in the non-tradable sector. Indeed, as the number of workers and equilibrium wages rise, so does the (local) demand of goods and services (e.g., hair-cuts, medical services, etc). In particular, the higher the level of human capital (skill-jobs), the larger will be these effects, especially for high-technology jobs. Clearly, the presence of direct "local" or "diffuse" multiplier effects hinges upon the location of where these additional goods and services are being spent (see, for example, Moretti, 2010 and Moretti and Thulin, 2013).[4]

For what concerns the role of public policies, like industry subsidies or tax-cuts, so far results have not been traduced into higher job-creation rates nor more innovative activities from firms (one of the major source business dynamism).[5] As a matter of fact, several stylized facts observed from the data suggest that the U.S economy is facing a slowdown.

On the one hand, some argue that the economy is running out of ideas despite outstanding efforts in a context where productivity rates are in continuing fall (Bloom et al., 2017). In that vein, the pace of fruitful innovations is reaching to an end and above all, many inventions that took place during the Second Industrial Revolution (e.g., transportation, communication, etc) could have happened only once (Gordon, 2012).

On the other hand, the rise of "superstar firms" along with industry concentration and stagnation of technology diffusion measured by the quantity of patent and citations (see, for example, Autor et al. (2017, 2020) and Akcigit and Ates, 2019*b*), could also be responsible for the decreasing job creation rates in the U.S economy. For instance, one firm or a small group of it—which are more productive— can take advantage of their position due to the access to new competitive platforms or simply by using high-quality goods with lower marginal costs like software production and online services. Thus, capturing a large market share with a relatively small workforce (e.g., Facebook or Google).

[4]According to Bartik and Sotherland (2019), such estimates only show a part of the agglomeration economies and congestion effects as they are calculated at the sample mean shares. In fact, the problem arises because Moretti uses the actual change in manufacturing as a projected instrument of total employment while at the same time, the variables of the model do not take into account the shock in terms of total employment. Hence, the multipliers reported in the second stage —depending upon the initial sector share— will tend to display higher values and vary across the sample. Consequently, state and local policymakers decisions could be biased towards the implementation of economic development programs to attract only high-tech firms.

[5]For instance, Trogen (1999) showed that Mercedes received an incentive tax-package of 165,000 usd per-worker to reallocate its plant into a different State (Alabama) but its cost was almost equal to 300 usd million dollar plant. Clearly, the ultimate goal of such policies is much broader than reducing fixed and variable costs like free land or industrial park sites. In fact, the combination of entrepreneurial policies (e.g., programs to promote to R&D, research on science conducted by universities, etc) are also important factors that may enhance regional labor markets and foremost avoid the lavish of scarce public funds.

In the same way, there has been a decline in the dispersion of new firms activities around the new millennium, particularly in the high-tech and information industries with large skewness-driven declines (Decker et al., 2016*a*,*b*).[6] Also, the global labor share has been declining significantly since the early 1980s —for different industries and across countries— due to the lower price of investment goods and capital augmenting technologies (see, for example, Karabarbounis and Neiman (2014) and Kehrig and Vincent (2018) for a manufacturing analysis).

In addition, markups for the high-tech and manufacturing industries have been rising during 1988-2015. However, the level of concentration in the former sector rose sharply (especially after 1995) with respect to the latter which remained more stable across the same period of analysis (see, for example, Hall (2018), Crouzet and Eberly (2018), De Loecker et al. (2020), among others). Likewise, decreasing investments, lack of technological opportunities and economic shift activities away from manufacturing (mostly to services) cope with the rise of "superstar firms" could have had an impact over labor productivity and firms' innovation activities. As a result, industries are now more sensitive to supply side labor policies. More precisely, there is a negative impact on labor productivity in industries where firm specificities and tacit knowledge are the main ingredient for innovative competencies. Hence, all knowledge tends to be concentrated in people, leading to high rates of labor turnover and making knowledge accumulation more difficult (Kleinknecht, 2020).

Nonetheless, in light of the current state of the U.S economy, is there still room for infrastructure investments? Nowadays, infrastructure and regional development are considered from a multidimensional perspective. In fact, not only geographical patterns or intangible assets matter, but also innovation policies as well. Indeed, the latter are important as they support the proliferation of new startups activities and at the same time, they encourage entrepreneurs to take advantage of new opportunities in the form of capital knowledge e.g., broadband investments (Audretsch et al., 2015).[7] Therefore, the implementation of large public infrastructure projects, could be regarded as labor enhancement policies for regional economies in terms of new jobs and better salaries and/or potential incubators for the development of high-tech industries like computer and electronic product manufacturing, telecommunications, ICT, among others.

[6] Even after controlling for firms' size, age and industry's type, these results only accounted no more than one-third of the decline in job reallocation (Decker et al., 2014).

[7] Innovation effects are not uniform because firms located nearby better transportation networks can absorb more resources from those investments than laggard ones located in distressed regions (Cohen and Levinthal, 1990).

Following the previous line, a recent work of Bennett (2019) takes an intermediate venue by analysing entrepreneurial dynamism and infrastructure provision (public and private) across U.S states. Surprisingly, he finds a negative relation between public infrastructure and business dynamism. In broader terms, job-destruction rates are positively affected by public provision whereas the converse occurs for private infrastructure and job-creation rates. Despite not regarding an industry analysis nor formally considering innovation activities, his results suggests that in the former case, public infrastructure works as an entrepreneurial "disabler" whereas in the latter, private infrastructure functions as an "enabler" of business entrepreneurship. For what concerns wages, a recent paper of Kemeny and Osman (2018) focuses on the differences between tech and non-tech jobs controlling for occupational skills and job characteristics (e.g., creativity, originality and schooling requirements). Yet, their empirical analysis disregards the effect of infrastructure investments and authors fail to find any systematic difference between different type of jobs. Therefore, these results can be considered a puzzle.

This paper fills the gap of the regional development literature by examining infrastructure investments (public and private), State incentives (monetary and non-monetary) and innovation outcomes (i.e. technological patents and scientific citations) in the U.S. One of the novelties of the present paper is the inclusion of non-monetary incentives (given by the diffusion of past public policies applied since the 1913's) and the production of technological patents conditional on industry subsectors value share to examine the performance of regional labor markets in terms of new jobs and better salaries. In empirical grounds, the paper aims to answer two questions: Can development incentives (e.g., R&D tax-credits) or else non-monetary incentives coupled with infrastructure investments (both public and private) boost job creation rates among industries? If so, how effective are these policies and foremost in which industries those effects will be larger? Furthermore, to what extent infrastructure, incentives (monetary and non-monetary) can influence the (conditional) wage distribution of (within) industries? To put the discussion into content and address those questions, I first identify the sectors for which infrastructure and innovation policies have the largest effects: high-tech industries like computer and electronic product manufacturing, telecommunications and ICT. To our understanding, its inclusion hinges upon two major reasons.

First, regional employment rates are expected to be benefited from high-tech firms as these industries increase regional aggregate value shares. Second, these group of industries can generate different forms of agglomeration economies and thick market externalities. Thus, its proliferation may lead to sizeable effects in terms of new jobs and salaries for regional labor markets (see, for example, Moretti and Wilson (2014, 2017) in the R&D and biotech industries). Then, I propose a simple regional model in which workers have some kind of mobility and firms make use of both type of incentives: monetary (e.g., tax-credits) and non-monetary (the mere diffusion of past public policies) in order to highlight the channels.

The first contribution of the paper is the identification of important stylized facts from the U.S business slowdown (e.g., higher knowledge gap and lower market competition between top and bottom innovators in public firms). That is, firms with lower marginal costs are more likely to be benefited from these public policies. The second contribution is an empirical assessment of the channels for which the above-mentioned policies work in regional labor markets. To this end, this paper employs Instrumental Variable (IV/GMM-HAC), Instrumental variable with arbitrary clustering (IV-AC), Spatial Lag Models (SLX) and Instrumental Quantile Regression (IQR) methods. I find that both types of incentives coupled with non-residential private spending have a positive business "enabler" entrepreneurship effect on employment rates in top value-added industries while public highway expenditures display a "disabler" effect only in sectors with lower value share. Interestingly, the cumulative response of generous state subsidies and R&D user cost are positively associated to higher job creation rates in bottom rather than top value-added sectors. However, when spatial variables are introduced into the analysis, I find positive (direct) effects of public infrastructure and insignificant ones for private infrastructure in total high-tech employment. Finally, interactions between non-monetary incentives and innovation outcomes in high-tech sectors with a higher aggregate value share are positively associated to employment networking (spillover) effects in other states provided that firms stimulate their local economies through increasing rates of ideas (i.e. production of technological patents) due to generous tax-credits. These trends are also observed in high-tech wages in sectors where marginal costs are already lower. That is monetary incentives imply additional cost-saving effects for high-tech firms and workers in those sectors have on average better salaries while a higher policy score does not imply better salaries. Therefore, the median worker in lower value-added sectors benefits more (on average) from these public policies as his/her salary is higher compared to those in top value-added sectors.

Overall, this is paper makes important contributions to the "jobs to people" literature as it links infrastructure, past public policies effectiveness and innovation outcomes in high-tech sectors where agglomeration economies and thick market externalities are considerably larger with respect to traditional industries related to these type of investments (e.g., construction). Thus, the empirical evidence in this paper is an important departure from the "beggar the neighbour effect" (see, for instance Wilson, 2009) because infrastructure investments couple with incentives both (monetary and non-monetary) indirectly generate positive spillover effects in other high-tech industries. From a state's perspective, efforts should be concentrated on the empowerment of laggard high-tech sectors, the reduction of the innovation/wage tech gap and investing R&D subsidies and federal grants into new low-cost technologies, job-training programs and/or dynamic processes based on worker's capabilities with soft skills. To the best of my knowledge, this paper is the first attempt to unpack employment and wages relations from infrastructure expenditures (public and private), development incentives and innovation outcomes (i.e., patents and scientific citations) at a regional level for industries with different aggregate values.

The rest of the paper proceeds as follows. Section 2 discusses the relevant literature. Section 3 outlines a simple model and the econometric approach. Section 4 discusses the empirical results. Finally, the paper concludes in section 5 and offers some policy implications.

2 Related Literature

Infrastructure investments coupled with the presence of local monopolies that restrict the flow of ideas to others, allow the internalization of innovation through technological waves. In plain words, entrepreneurship entails innovation because the essence of capitalism is to evolve i.e. to seek new combinations —of products, production processes, new forms of organizations and economic opportunities— while destroying the existing ones in a creative way while reallocating these resources into the highest valued use (Schumpeter, 1934, 1942). Others instead consider that the most important transference of knowledge comes from the outside (Jacobs, 1969). That is public policies like infrastructure expenditures provide the ingredients to boost regional convergence (Armstrong, 1995) through increasing rates of innovation activities among firms, new establishments, jobs and/or better-paid salaries for local economies.

In the case of infrastructure, whether we regard the public: "broad" (health, justice, education, transportation and facilities), "core" (highways and bridges, air transportation, sewage systems) or the private[8] definition (e.g., non-residential spending), they all support and facilitate diverse activities performed by workers and firms (Miller, 2013). Indeed, these traditional expenditures not only stimulate the proliferation of new establishments, but also they may produce sizeable effects through increasing rates of innovation, changes in the wage structure among industries, among other factors.

Nevertheless, at a macro-regional level the effects of infrastructure investments over labor markets are mixed. For instance, considering a universe of 12 industries, Pereira and Andraz (2003) find that highway expenditures can shift the sectoral composition of employment in the construction, durable manufacturing and transportation industries. Conversely, private investments may foster the overall pace of employment reallocation in manufacturing, public utilities and communications sectors. In broader terms, public infrastructure investments affect private employment positively in only six of the twelve industries under consideration while private investments in only five. Furthermore, public investments tend to be absorbed disproportionally more in bigger states (Pereira and Andraz, 2004, 2012*a*,*b*) and may contribute to expansion of regional asymmetries —in terms of employment growth— because the major source of funds, hinges upon the number of representatives in the senate as well as the extent of (previously) approved projects.[9]

Moreover, the work of Leduc and Wilson (2013) explores the dynamic effects of highway spending through the lens of federal grants. Yet, given that economic agents are foresighted (i.e., they may anticipate well in advanced the eventual level of grants); as a result, capital outlays tend to be slack i.e. endogeneity is likely to occur. In order to overcome this issue, authors take the expected present value of forecasts (current and future) grants as a source of exogenous variations. Overall, their impulse response analysis suggests no evidence of an initial impact over employment, unemployment or wages. Although, for some industries like transportation and retail the effects were large in the medium run (6 to 8 years out). Likewise, to account for non-linearities, they assess the validity of predicted grants for periods in which the GDP contracts or expands (i.e. recessions and expansions). Results are imprecise, hardly any difference after 8 years can be observed, though the initial impact of the former variable

[8]Even though infrastructure is a non-rival good subject to congestion, private firms usually invest on these projects as a mean to: 1) reduce delivery times; 2) improve their plant (capacity) in the manufacturing of goods and services and 3) decrease their production costs (technical efficiency).

[9]The identification strategy subsection exploits this fact by using the number of representatives in the appropriation committee as an exogenous variation source in highway expenditure.

over employment is indeed significant. Interestingly, a recent work of Greenaway-McGrevy and Hood (2019) refines the Vector Autoregressive approach (VAR) by filtering-out the effects of common shocks like national crisis over U.S regional economies. Results show that after a shock, employment recovers in the following period. Therefore, filtering methods provide more flexibility —with respect to the Blanchard and Katz (1992) canonical approach—, to capture the spatial temporal heterogeneity of labor market indicators without including specific controls.[10]

In spite of the inconclusive evidence, infrastructure investments are paramount as they effectively influence firms' location by releasing additional resources, favouring knowledge accumulation (e.g., patents and scientific citations) and generating positive externalities in the regions where those ideas were originally developed (see, for example, Jaffe (1989), Rauch (1993) and Feldman and Florida, 1994). Nevertheless, regional asymmetries can be persistent due to lower capital stocks (physical and human), differences on the distribution of geographic resources (e.g., land fertility, temperatures) and heterogeneous demographic characteristics like a higher proportion of qualified workers on richer regions (see Capello and Nijkamp, 2010, chap. 2 and chap. 10).

In the same way, the knowledge spillover theory suggests than richer regions are more likely to generate more entrepreneurial opportunities than poorer ones due to the context in which entrepreneurs grasp the decision-making process (Acs et al., 2013). For example, the higher the value of regional amenities offered to workers (e.g., housing quality: more space, access to fireplaces, better water systems, etc), the more concentrated will be the creation of knowledge as (the share of) high-skilled workers will choose to live in the nearby of metropolitan areas. Hence, fostering job-creation rates among firms but also influencing the wage-premia across regions (Moretti, 2013).[11]

For what concerns incentives and funding sources, taxes reflect (at least partially) the opportunity cost of infrastructure spending. In addition, an excessive fiscal burden can increase or reduce land value through regional amenities (e.g., better highways, sewage systems,etc) and foremost influence innovation activities performed by firms. Therefore, taxation affects the location of both workers and firms giving rise to several effects over regional economies. In that

[10] However, an important drawback of this approach is that both the number of units N (cross sections) and T (number of time series), must be sufficiently large to consistently estimate all components of the model, a requirement that may not be fulfilled.

[11] Of course, due to the unobserved ability among college workers, this wage-gap could be biased downwards or upwards although these differences are generally orthogonal from housing costs.

vein, the empirical evidence is also mixed.

Studies show that State incentives or "jobs-to-people" policies in the form of: investment credits, R&D stimulus, property and job-creation tax-rates, are far more efficient than "people-to-jobs" (i.e., out-migration subsidies) because encouraging people to leave those regions not only reduces the demand (i.e., through less households spending), but also it produces a negative effect on the supply side through lower housing prices and decreasing construction rates. Thus, as population densities decrease, more jobs will be destroyed generating a negative impact over individual's wealth. For example, ever since the 1990's, incentives have tripled and they have been intensified (e.g., Wisconsin's Foxconn manufacturing $3 billion plant) and yet employment rates have remained stable across time (Trogen, 1999). In that same line, Garin (2019) find that the ARRA event study program effects were slightly positive and close to zero at an aggregated level. For instance, in the short-run for each dollar spent on the construction sector in the year 2010, six jobs were created per $1 million.[12]

Studies empathize that generous tax-credits have a positive but limited effect over sectoral employment and innovation. For instance, focusing on the U.S biotechnology industry at both margins[13] (intensive and extensive), Moretti and Wilson (2014) show that development incentives through R&D subsidies raise the number of star scientists. However, most of these gains are the consequence of state-adopters relocations with limited effects on incumbent scientists productivity. In terms of total employment, they uncover significant (positive) gains in the Pharmaceutical and Medical sector (16%); the Pharmaceutical Preparation Manufacturing (31%) and the scientific R&D industry (18%) and almost no effect or close to zero for salaries in those sectors. However, despite considering tradable and non-tradable sectors outside the biotech (e.g., high-tech industries like chemical, machinery manufacturing), the inclusion monetary incentives are insignificant different from zero or negative, suggesting no stimulative effects over employment. Surprisingly, efficiency gains on the R&D user cost were insignificant in all cases. A

[12] According to Ramey (2020), there were at least two facts for which the ARRA was not effective. First, at that time, interest rates were closely to zero (lower zero bound). Second, unemployment rates were on average 9% to 10%. Hence, it appears that the monetary (accommodative) policy applied by the FED may have wiped out the potential benefits of the program leaving slightly positive or either negative effects.

[13] This paper main focus is based on the intensive response (employment growth) to specific shocks: highway expenditure and development incentives (monetary and non-monetary) as a proxy of innovation policies. For what concerns the other margin, empirical evidence suggests that most of the variation comes from the extensive entry margin, not from the behavior of incumbents. In other words, a local shock increase wages, attracts more workers to the area spurring the creation of new entrants. Hence, the created externalities propagate an amplify the process of firm creation by reducing exit rates. The latter pattern is also captured when firms are large, have fat tails and perform external innovation activities (see, for example, Walsh (2019) and Cao et al., 2020).

possible explanation offered by the authors is that scientists in those sectors represent a smaller share of total employment compared to the biotech industries or possibly their labor supply is more inelastic.

Interestingly, new evidence shows that monetary stimulus give residents extra-job experience and can influence regional income distribution, generating positive spillover benefits even in the reduction of crime and child development (Bartik, 2019*a*). However, they may also lead to fiscal externalities by inducing firms and workers to reallocate. These decisions, impact the demand and supply of labor, which ultimately hit population, income distribution and the labor market. In this line, Bartik (2019*b*) points out at least three reasons for which those policies have not been effective at stimulating regional and local labor markets.

First, they are not target-oriented to distressed places (e.g., Indiana and Illinois have the same employment rates but the former receives twice of the economic incentives). Second, the extent and the scope of incentives is extensive are fairly dispersed to all kinds of firms. Indeed, the average duration of incentives is about 10 to 15 years —a period beyond business owners and young firm survival rates— and foremost, they are not focus on the major source of employment growth (i.e., high-tech industries).[14] Third, public policies that are not aimed at improving firms' inputs (e.g., updating roads or other public infrastructure services) can spur job-creation rates and ultimately offset the full potential benefits of development incentives. In fact, given the complementary between firms' inputs and infrastructure expenditures, improvements in productivity may boost even more regional economies if tailored-business services have a considerable weight on labor markets. For instance, manufacturing extension services or customized job-training programs by community colleges supported by federal aid, can provide advice to smaller manufacturers on adopting new low-cost technologies, target-specific market research or better soft-skills for profitable business networking. Accordingly, the combination of both tax-incentives and infrastructure expenditures might lead to increasing returns of employment growth if the workforce is highly skilled and trained. Certainly, as the number of available jobs augment, state policymakers could make use of those additional resources for more job-training programs.

[14] Another important drawback of development incentives is that current schemes are biased towards large firms rather than small ones. In that way, downstream employment creation is unlikely to be successful because for the latter type of firms, business reallocation becomes more costly. Hence, in order to improve tax-incentives effectiveness, federal intervention should be considered.

On the other side, tax evidence shows that narrower tax-bases not only discourage firms' entrance, but also reduce the expected net profits from specific locations (e.g., a higher cost of living), employment rates and real wages (Serrato and Zidar, 2018). Albeit, the pace at which firms enter onto the market have important macroeconomic consequences, the effects of a smaller tax base over employment rates and real wages will ultimately depend on the state income incidence by workers, landowners or firms owners.[15] For those reasons, tax-cuts towards lower or moderate income groups within a spatial discontinuity framework instead of taxing the top 10% "wealthiest" decile are more effective to increase employment rates (see, for example, Ljungqvist and Smolyansky (2014), Zidar (2019), among others).

For the same token, taxation produces a negative effect over innovation outcomes and may lead to human capital reallocations. In this line, Akcigit et al. (2018) show that the elasticity of innovation in the U.S (measured by patents and citations) with respect to both taxes (personal and corporate) at a macro (State) level is higher for workers inputs (expenses and effort) as a result both factors become more costly. Accordingly, inventors will decide how much effort to assign to their own work and companies will choose the amount of qualified workers to hire. For those reasons, each state might have a different response to infrastructure provision due to citizens/firms heterogeneity in regards to the potential benefits in different states and/or bordering counties. For instance, high-productivity inventors could "shirk" at their workplace because shared profits with firms are not attractive. Therefore, regardless the tax-base provided by each state, firms would be willing to offer lower wages, or else inventors might abandon firms in the search of more profitable opportunities in other regions or more precisely, the so called "business stealing" effect.

In spite of the negative effects of taxation, policies should be targeted at the most innovative industries.[16] Earlier studies at a micro level (see, for example Haltiwanger et al. (2014, 2016),

[15]For instance, Amior and Manning (2018) claim that employment rate can be expressed as a "sufficient statistic" i.e. all workers are renters, so housing costs can only affect welfare through real consumption wage or a local projection mix of employment. However, if residents are also owner-occupiers, then, housing becomes mobile and changes in housing prices may curb population after controlling for employment rate. Clearly, the sign of this effect is ambiguous, but as authors explain, the inclusion of that variable does not seem to be statistically significant.

[16]The definition of which sector belongs to the high-tech category is not exempt from drawbacks. In fact, according to the BLS, an industry corresponds to the high-tech category only if its employment share exceeds twice the overall national average (with the special focus on science, engineering, and technician occupations). Yet, as firms gain market power, labor shares decrease and output falls; thus, real wages will follow that trend because prices are higher (De Loecker et al., 2020). Therefore, the share of industries varies across time because highly qualified workers may abandon firms in the search of more profitable opportunities. As a result, since one of our

among many others) documented that after the 2000 period, young high-tech firms were responsible for more than 20% of new jobs. However, their rate of survival is not long enough (5 years or less) and above all, the rate of created jobs decreases monotonically with firm size.

Others instead argue that high-tech firms generate more profitable opportunities for entrepreneurship and employment growth because it is less expensive and less time consuming for geographically separate individuals and firms to acquire new learning skills and come up with the development of brand new technologies. Hence, regions in which institutions have a better performance, are more likely to provide high-quality public goods (e.g., highways), affect entrepreneurship dynamics (i.e. innovation spillovers) through governance indicators and better access to capital (see, for example, Williamson (1998), Acs et al. (2008), among others). More recently, Tuszynski and Stansel (2018) find a positive and significant relationship between economic freedom and entrepreneurial activity in the U.S (measured as granted patents per 100,000 population) but a negative robust relationship between patents activity and the sum of economic incentives: property tax, job-training grants, job-creation, tax-credit and research & development when establishments are "small" (less than 10 employees) while the converse is observed for "large" firms (i.e. those with more than 500 employees). Accordingly, the institutional environment in which entrepreneurs interact among each other curbs innovation incentives. That is lesser regulations in the labor market coupled with a reduced government size in terms of taxes and spending and development incentives, strengthens business entrepreneurship in the short-run by enhancing job-creation rates with mixed effects over welfare depending on establishment size.

On top of the above-mentioned evidence, Bartik and Sotherland (2019) claim that regional employment is expected to be benefited from high-tech firms because these firms can generate different forms of agglomeration economies. Thus, the larger the place and/or concentration levels of these type of industries, the higher will be the spillovers and employment clusters. Moreover, high-tech firms may contribute even more to the development of laggard labor markets if the initial employment-share is around 10-15%. In fact, there is a threshold effect beyond the upper bound (15%) in which the high-tech multiplier is modest or it does not lead to employment

concerns is related to the distribution of innovation in terms of aggregate value, we take a standard approach in considering all industries within the software and telecommunication sectors. To this end, we follow the strategy suggested by Hall (2018). In broader terms, our choice is not that far from Bartik's new labor shares estimates (the Appendix B of the aforementioned paper lists the new ranking of high-tech industries).

growth. Therefore, agglomeration economies may have diminishing returns (i.e., a sort of an inverted u-shape) or congestion effects could deter employment growth. Or else, it could imply the presence of gap between the arrival of a shock and firms' response in regards to agglomeration economies that foster productivity and wages.

For the same reasons mentioned above, industry wages should not be disregarded from the analysis as high-tech sectors are constantly evolving over time. Certainly, infrastructure investments can be regarded as a sort of "amenity" (e.g., reducing time distances) for all type of industries (including the high-tech) as it is a non-excludable good. Earlier studies show that infrastructure expenditures benefit both manufacturing and retail industries attracting additional workers and in-migrants towards more developed regions, boosting employment rates and foremost increasing real wages without offsetting private sector productivity (Dalenberg and Partridge, 1997).

Furthermore, new evidence suggest that even though young high-tech firms bring employment and innovation gains over labor markets, these effects tend to vanish as firm grow in size and age (Santoleri et al., 2019). In other words, there is a positive (though heterogeneous) effect of innovation activities (R&D expenses and granted patents) over the conditional distribution of employment. Nonetheless, throughout quantiles employment rates increase during the first years but as firm grow old, they tend to diminish. In the case of patents activity, employment growth is positive and statistically significant starting from the 30^{th} up to the 90^{th} percentile with a particularly larger effect over newborn firms. Therefore, the solely addition of new establishments and/or an increase in the pace of innovative activities, do not necessarily imply more jobs nor higher real wages. Interestingly, Kemeny and Osman (2018) explore the wage gap between tech and non-tech activities controlling for a set of occupational skills and job characteristics (e.g., creativity, originality and schooling requirements). Results show that the effect of tech employment on real wages does not vary across different levels of characteristics. In addition, despite restricting the analysis to very high-paying jobs (e.g., finance and insurance), the difference between tech and non-tech remains unaltered which could be considered a puzzle.

For those reasons, the present paper provides an alternative explanation and argues that an uneven distribution of aggregate value share within high-tech sectors could explain such puzzle. In plain words, as we will later show, there is an heterogeneous effect over the conditional distribution of wages due to monetary and non-monetary incentives (for more details see subsection

3.2). Instead of assessing the effect of individual policies, we hypothesize that infrastructure expenditures, R&D tax-credits, the production of scientific knowledge (conditional on aggregate value share) and the diffusion of public policies may explain the asymmetries of high-tech firms (i.e., small and large) in terms of aggregate value share and foremost in which sectors the effects over regional labor markets (employment and wages) will be higher.

Accordingly, in order to identify the casual response of regional labor markets to different shocks, an econometric model will be presented in the next section. Broadly speaking, infrastructure investments (public and private); monetary incentives (R&D tax credits); non-monetary ones like the mere diffusion of public policies in the fields of education, labor markets, domestic commerce; and innovation outcomes measured by technological patents and scientific citations conditional upon firms' aggregate value share may be influencing regional labor markets through several and (possibly) opposed channels: 1) geographical (e.g., temperatures and regional characteristics); 2) socioeconomic (e.g., high-school and college graduation rates); 3) institutional quality (e.g., labor market friction index) and 4) supply shocks (e.g., recessions). As a result, employment rates could shift upwards or downwards depending on the extent of each channel. Thus, we assume that workers have some kind of mobility. Yet, highway outlays can be correlated to employment measures (i.e., endogeneity). To this end, in lieu with extant empirical literature, we appropriately address this issue by regarding the number of senators in the committee members as a source of quasi-exogenous variation for infrastructure spending (further details are in the identification strategy subsection 3.2).

To conclude our empirical analysis, we exploit high-tech heterogeneity to assess the possibility that the above-mentioned channels may have a different response at part of the wage distribution. Simply put, one may wonder to what extent conditional quantile wages are affected by changes in state policy interventions, like infrastructure investments and per-capita patents relative to industry's value share and/or the mere diffusion of past policies. Overall, the aim of this exercise is to identify the most profitable sectors in which those policies have the largest impact over regional labor markets.[17]

[17] Of course, there are many other factors that certainly affect the wage distribution. For example, the presence of stringent regulations (e.g., minimum wages , workers' licenses), degree of (de)unionisation, workers' occupation, industry market power, gender inequality, etc. A comprehensive study of infrastructure expenditures considering all those aspects is beyond the scope of the current paper.

3 Econometric Approach

Our empirical strategy relies on the workhorse of urban economics studies Roback (1982) and Gyourko and Tracy (1989). In this context, we focus on the effect of both monetary and non-monetary incentives over regional (state-level) labor markets for high-tech firms. The research questions are tackled in the spirit of "jobs-to-people" policies in which local policymakers stimulate their economies through infrastructure expenditures (public and private) as well as generous subsidies incentives (R&D tax credits) towards firms. In this vein, we regard all U.S public firms related to the software and telecommunication industries (high-tech).

One of the novelties of our approach is the inclusion of non-monetary incentives proxied by a score on the diffusion of past public policies in several fields (e.g., education, energy, environment, domestic commerce, etc) as *innovation boosters*. In broader terms, the introduction of both type of incentives not only generates *direct effects* over regional labor markets (e.g., more jobs and payrolls), but also it may increase the production of scientific knowledge (i.e., technological patents) conditional upon the share of aggregate value within the industry. For instance, sectors with lower marginal costs like software production and/or online services are more likely to generate more knowledge clusters.

For what concerns the *indirect effects*, we are aware that agglomeration economies (e.g., higher local multipliers due to incentives) outside high-tech sectors as well as different responses at the extensive margin could be at work (e.g., new birth establishments). In this line, as explained earlier, the aim of this paper is to assess the cumulative response of infrastructure and innovation policies only at the intensive margin. Therefore, we will not pursue that road when performing the econometric analysis.

Having mentioned the facts from above, before proceeding to the econometric model, a simple model is proposed in Figure 1. For the sake of the brevity, we leave the formal proof outside but we briefly explain its intuition. Let us consider a world in which workers maximize their utility from a composite good (Q), land (N), a vector of natural amenities (A), infrastructure services (G) with an inelastic labor supply (for the ease of the analysis). At the same time, firms maximize their utility (choosing labor) given wages (W), taxes (τ), a vector of local amenities (A), making use of infrastructure services like highways (G) with free entry/exit (long-run) which implies profits equal zero. In addition, if we assume additive separability and conditional independence with forward looking decisions (Bayer et al., 2016) and solving for employment taking as given a set of past location characteristics, then it is possible to disentangle the direct and indirect effects of both types of incentives (monetary and non-monetary).[18] Accordingly, we will have the following effects onto labor markets.

First, the adoption of monetary incentives in the form of R&D tax-credits towards local markets directly increases the size of the industry as it curbs inventors incentives to remain in the area in which those benefits are granted. Yet, as explained before, there is an heterogeneous and higher response of innovation elasticities to taxes at a macro level (Akcigit et al., 2018). Thus, despite offering positive incentives some skill-workers may be tempted abandon the region —possibly in bordering counties— in the search of more profitable opportunities. Hence, employment effects could be partially offset by this fact. In practice, regardless the industry under consideration, innovation is usually measured as the number of star scientists at a state-level (i.e., those in the upper distribution). In our case, we consider the conditional distribution of patents and scientific citations —in terms of industry aggregate value share— at each percentile under both types of incentives: monetary and non-monetary.

Strictly speaking, our framework includes R&D tax-credits, firms user cost, infrastructure services (both public and private) as well as the mere diffusion of past policies and/or per capita patents as innovation boosters for regional labor markets. In the case of public infrastructure a public good subject to congestion, new expenditures support and facilitate diverse activities performed by firms as these imply a reduction in labor costs. Hence, additional resources (a priori) could be assigned to more research activities and job training programs. Nonetheless, the creation of knowledge takes time as it cannot be diffused rapidly. Put differently, delayed and cumulative effects are more likely to be present as the potential benefits of lets say a new highway may fully be visible in the medium or long-run (a minimum of 5 to 10/15 years). As a result, we expect displacement (mostly negative effects in the short-run)[19] cope with positive ones (human capital, improvements on the quality of institutions and/or spillover effects) in the

[18] Note that we have made the strong assumption that local governments have an objective function and choose to subsidize those firms for which their potential aggregate value share may provide future benefits to their local economies. Since high-tech incentives impact all firms user cost, only local representatives can decide the implementation of tax-cuts of fiscal benefits for their economies. Hence, we might separate public firms in terms of aggregate value before the implementation of any incentive program.

[19] Another channel that produces a negative shift in local markets is out-migration or people-to-jobs subsidies. As we stated before, encouraging people to leave regions not only reduces its demand spending, but also it produces a negative effect on housing prices. Hence, as population densities decrease, more jobs are destroyed generating a negative impact on regional income. Likewise, the same effect is at work when a contraction in the national economy takes place.

medium/long-run. But in broader terms, we can expect an entrepreneurship enabler effect of private infrastructure as it can also facilitate the development of diverse activities and foremost generate cost saving effects.

Using our simply model we can observe that both types of incentives (monetary and non-monetary) may coexist simultaneously in the sense that they can stimulate innovation activities and increase aggregate value of the industry. In this vein, private capital investment in any form is the bridge between jobs creation and better salaries which is also consistent with targeted policies to distressed places or jobs-to-people that lead to an increased demand for goods and services.

Last but not least, it is noteworthy that our conceptual framework uses arrow dots to empathize conditional relations which in our case hinge upon each sector contribution (in terms of aggregate value share) from the national economy. Clearly, at a regional level it will depend on the interaction of several and possibly opposed channels. Along these lines, our theoretical road-map assigns indirect effects in regards to innovation and employment outcomes. To this end, high-tech value share is the vehicle for which different subsectors (with different marginal costs) create more jobs.

Finally, the approach employed in this paper relies on the large body of existing evidence (e.g., Moretti and Wilson, 2014). Certainly, it makes sense to have some kind of mobility among highly educated workers (i.e. elastic labor supply). However, this degree of mobility may not have to be the identical within each subsector. As a result, we expect an heterogeneous conditional response on payrolls based on incentives and their corresponding value share within high-tech industries. In other words, infrastructure and innovation policies affect the entire *conditional distribution* of industry wages because each sector embodies workers with different qualifications and also the production of knowledge is not uniform (i.e., some states may produce more knowledge than others). As a result, the contribution of each subsector in terms of aggregate value is also different. For those reasons, our second empirical exercise examines the impact of high-tech wages to infrastructure and incentives (monetary and non-monetary) and innovation outcomes through the lens of quantile regression models (further details of the empirical strategy are discussed in subsection 4.3.1).

Employment Models

Our benchmark specification departs from the standard panel-fixed effects model. We first uncover causal relations regarding both types of infrastructure and non-monetary incentives as a proxy of innovation policies for the 48 contiguous states using different controls taken from the urban economics literature. Formally, we test the following equation:

$$y_{it} = \chi_{it} + \iota_{it} + \kappa_{it} + z_{it} + f_i + d_t + \epsilon_{it} \tag{1}$$

Where y_{it} is the outcome of interest: job creation rates for the top and bottom value-added high-tech industries; χ_{it} is an infrastructure vector (highway and non-residential expenditures); ι_{it} represents the non-monetary incentives (i.e., the diffusion score of past policies net of infrastructure) and/or the conditional percentiles of per capita patents; κ_{it} is Real GDP growth rate excluding both measures of infrastructure (public and private); z_{it} a vector of urban controls named average temperature, human capital, institutions and supply shocks (i.e., economic downturns); f_i are division fixed effects and ϵ_{it} is the error component term. Additionally, since highway capital outlays vary in sample size, I include a dummy variable d_t to account for time effects. In this way, the stability of the panel given its short dimension ($N = 48$ and $T = 26$) is preserved.

Nonetheless, equation 1 does not capture the entire picture of our first research question. Consequently, we move on to a full specification in which we simultaneously test for employment effects under both types of incentives: R&D tax-credits; firm's user cost; non-monetary incentives (the mere diffusion a set of past policies) and/or the conditional percentiles of per capita patents along with the usual controls from urban economics models. Formally, our second equation is the following

$$y_{it} = \chi_{it} + \iota_{it} + \sigma_t + \gamma r_{it} + \kappa_{it} + z_{it} + f_i + d_t + \epsilon_{it} \tag{2}$$

Where the high-tech R&D tax-credit (σ) is a dummy variable equal to 1 for those adopters with positive incentives rates and 0 in contrary case. In this case, we adopt a more standard approach since we are also dealing with a common consumption good (i.e., highway). Further, we include firms user cost from U.S public firms (γr_{it}). Unlike previous studies, when measuring this variable we consider a longer time span (1965-2015) of the R&D expenditure shares which are on average half of the IRS (0.5). Moreover, since we are also interested on delayed effects, we directly allow for the cumulative response of user cost and we interact this variable with the dummy incentive to get a magnitude of monetary incentives over regional labor markets. The expected sign of the interaction is negative because when reduction in the R&D user cost takes place, it stimulates the number of star scientists at any level (corporate, private, academic) through higher saving costs; hence, realising additional resources for hiring qualified workers. Conversely, a positive user cost implies less R&D investments or less resources to investment in human capital; thereby, higher displacement effects and/or fewer star scientists amongst states. In addition, given that public capital outlays tend to be slack and incentives could be contemporaneously correlated, we allow delayed effects (up to 3 years) of the dummy variable.[20] Hence, we estimate a variant of equation 2

$$y_{it} = \chi_{it} + \iota_{it} + \sigma_{t-3} + \gamma r_{it-3} + \kappa_{it} + z_{it} + \sigma_{t-3} * \gamma r_{it-3} + f_i + d_t + \epsilon_{it} \tag{3}$$

Accordingly, our preferred specification will be equation 3. Therefore, if high-tech incentives have a positive effect on employment rates, then, $\sigma > 0$ and $\gamma < 0$. However, there are several aspects that are worth mentioning.

First, for what concerns firm's user cost, one might be sceptical about its variability and/or the degree of exogeneity with respect to the dependent variable. Nonetheless, as Moretti and Wilson (2014) claim, within each state political idiosyncratic factors as well as the implementation of incentives are not uniform. In terms of our research question, we assume that each unit can be considered independent (exogenous), but we also acknowledge the fact that development incentives hinge upon the political power of the different committees in the house of senators (e.g., small business and entrepreneurship, finance, etc) of each region. Another potential issue would arise if general incentives are correlated to unobserved differences across states for high-tech industries. If so, then, depending on the extent of such correlation, permanent differences might not be captured through fixed-effects and as result we would have a positive (negative) bias which may lead to overestimate (underestimate) the cumulated R&D user cost. Nonetheless, since we are in the presence of common consumption good (i.e., infrastructure),

[20] As a matter of fact, the sum of the coefficients on current and lagged incentives yields a cumulative effect of four years. Naturally, medium run effects in regards infrastructure imply the use of more lags. But given the short dimension of our panel, our econometric models follow the empirical literature and allow up to 5 lags of highway expenditures.

then, we assume that unobserved heterogeneity comes from the capacity of each firm (belonging to their corresponding subsector) to absorb infrastructure services and to re-invest those additional resources into R&D programs and hiring more workers.

Second, we introduce the notion of non-monetary incentives as a proxy for innovation policies which is novelty with respect to existing literature. Policy diffusion data can be fairly representative of state policy activity in different fields. Particularly, its use may provide more answers about the performance of regional labor markets. In our case, since we regard a set policies (excluding transportation) back from 1913's, it is likely that the smooth indicator be weakly correlated predictor to our labor market outcome. Hence, it enters into our model as an exogenous variable.

Third, as regards the effects of regional innovation, it is most likely that a positive sign on γ can be related to firm's dimension and/or an unequal distribution of their innovation activities due positive delayed effects of infrastructure investments (e.g., a higher number of senators might have led to more discretionary funds to most innovative states). In this line, we overcome this issue by accounting for the conditional contribution of each sector to the economy as a result of different absorbing capacities of infrastructure. More precisely, we weight patents distribution by their respective aggregate value share of the sector. In this way, we disentangle those sectors as the top value-added from bottom value-added.[21]

Fourth, an important fact that needs to be mentioned is the correlation among residuals and cross-sectional dependence among units. It is well known that persistence among predictors in regards to high-tech employment are likely to occur. Moreover, OLS estimates are naturally biased due to unobserved heterogeneity and cross-correlation. As a result, we account for both issues within an instrumental variable (IV) framework data following two different roads.

On the one side, we first estimate equation 3 using the HAC (autocorrelated and heterokedastic) correction. Then, as a first robustness control we make no assumptions about the spatial structure of the error term. In plain words, correlation within each units takes place in any possible way, without any kind of imposed structure (Colella et al., 2019). This is possible

[21] Another strategy would be to regard deep historical innovations as a weakly correlated predictor of (current) high-tech wages due to different levels of the technological frontier (Petralia et al., 2016). However, this type instrument is time-invariant and foremost to be informative, it requires that both cross-sectional and panel estimates be sufficiently large to be compared. In our case, we focus on innovation outcomes from a small group of U.S high-tech public companies whereas historical corporate patents could be classified under several industry codes (i.e., one firm could be registered under several activities). Hence, since infrastructure expenditures is one of our major shocks, we simply use its past lags as a predictor to account real wage effects.

because unlike traditional spatial frameworks, groups are non-overlapping in each dimension and each observation belongs to the corresponding group. In this way, we avoid the problem of serial correlation in both dimensions (state and time) which might create mismeasurement in standard errors. Unlike the work of Moretti and Wilson (2014) in which cumulative effect are defined as the ratio of the coefficient over the pre-adoption mean (holding constant R&D tax credits), our framework employs further lags because infrastructure full benefits are more likely to occur within a period of 5 to 10 years. Thus, by assuming a flexible relation among errors, serial correlation can be ameliorated.

On the other, we estimate a Spatial Lag Model (henceforth SLX) while accounting for cross-sectional dependence (CSD). It is well known that CSD is an issue present in almost all macroeconomic datasets because of the nature of aggregate data. In this vein, depending on the DGP, the presence of common, time varying relationships, etc, several strategies can be considered (see, for instance Pesaran (2006), Gunnella et al. (2015), Bailey et al. (2016) among many others). In the case of short unbalanced panels like the present paper, the above mentioned strategies do not apply given that the number of units and time span should be relatively large. Literature provides evidence that filtering methods are consistent in the presence of both weakly and strongly correlated panels (Greenaway-McGrevy et al., 2012). Accordingly, a simple defactoring procedure may be sufficient to reduce the extent of common factors and strong CSD among units. Following Bai-Ng filtering methods (Bai and Ng (2002), Bai (2004) and Bai and Ng, 2004), the BN criteria is applied to the (standardized) first difference dependent variable (high-tech employment) provided that the DGP may be explained by an AR(1) structure process (which applies to this DGP) yielding 3 factors according to the IC2 criteria.[22] However, it is worth stressing that first-difference filtering requires no gaps on the variable of interest. For that reason, the empirical results of the SLX model are based on the overall response of high-tech employment and not by our value-added classification. Similarly, the inclusion of conditional percentiles are also subjected to the same problem. Therefore, dummy variables for different industries are included to support the stylized facts and the previous empirical (non-spatial) results.

[22] If the process does not exhibit strong serial correlation, then, the LSDV filtering outperforms the first-difference approach.

3.1 Data Description

When exploring infrastructure and innovation outcomes over private employment for the high-tech industry, this study combines different datasets. Our main shock (public infrastructure) is constructed using the capital outlays of highway spending which includes the following items: maintenance, operation, purchases of equipment, toll highways, bridges, tunnels, ferries, street lighting, snow and ice removal (Pierson et al., 2015). In doing so, we follow other studies (Hooper et al., 2018) that make use of the same specific expenditure. Likewise, we are also interested on private infrastructure responses over labor markets. Hence, we include non-residential spending as a proxy of private infrastructure. Overall, our infrastructure measures (public and private) rang from 1990 to 2015 although the latter variable is only available from 1993 onwards.

An important task is to compute real measures. A usual practise is to divide nominal expenditures by a state price index provided by the Bureau of Economic Analysis (BEA) which takes into account the price of investment goods (for more details see table 3.9.4). Conversely, we deflate private infrastructure using the Bureau Labor Statistics (BLS) consumer price index (CPI-U 2012=100). What is more, in order to control for the business cycle effect, both spending variables are defined in per capita and later transformed into logarithms (Dupor, 2017).

For what concerns innovation, we rely on two measures. Firstly, we concentrate on the conditional distribution of innovations from high-tech industries. In other words, we take the weighted percentiles of the entire distribution of patents and scientific citations using the share of aggregate value of U.S public firms in 2001 (Crouzet and Eberly, 2018, Hall, 2018). In lieu with economic literature as well as the U.S. Patent and Trademark Office's (USPTO), we rescale patents per 100,000 population. Patents data comes from patents view, a crosswalk of U.S public firms (COMPUSTAT) Dorn et al. (2020) and complemented with Arora et al. (2017). As regards the definition of high-tech sectors, we regard a more traditional approach and consider all software and telecommunication sectors at three-digit NAICS and KLEMS levels. However, a few limitations of our innovation measure are worth discussing. First, our universe of public firms represents a small fraction of the high-tech industry from national economy. In our case, we picked-up patents and scientific citations from 1,439 public firms which perform several activities related to our high-tech classification. As a result, our empirical estimations will regard the total number of patents and scientific citations from the aggregate sector and not by subsectors due to the lack of observations. Second, in order to have a clear picture of the industry in response to different public policies, it would be desirable to have information from both types of firms (public and private). In that line, the work of Dorn et al. (2020) presents a promising avenue for future research on the drivers of private innovation.

A dynamic innovation and policy diffusion score across U.S states of different public policies applied since the beginning of the 1913's is computed (Boehmke et al., 2020). We construct a smooth index for each state depending on the type of policies adopted. We focus on macroeconomics, labor, education, environment, energy, housing and domestic commerce policies leaving aside transportation since our goal is to highlight the interaction of monetary, non-monetary incentives and innovation outcomes over regional labor markets. Overall, our reduced policy diffusion score encompass a total of 204 policies (details of each policy are available upon request).

As regards private employment and industry wages, our preferable source is the Quarterly Census of Employment and Wages (QCEW). Data is based on administrative records (state Unemployment Insurance payroll reports) covering all employers with no minimum thresholds and/or flag data for employer size. Hence, they contain minimal measurement errors. Unfortunately for the latter, the universe of firms is narrow. Accordingly, we complement the payroll from County Business Patterns (CBP) as it provides more information. However, data requires an adjustment because it was measured under different industry classifications between 1990 and 2015. Therefore, we take the weights from Eckert et al. (2020) and adjust all SIC values into NAICS 2012 levels.

For what concerns monetary incentives, we follow the existing literature on tax-incentives and employment by taking R&D tax-credits from the Panel Database of Incentives and Taxes (PDIT) (Bartik, 2017).[23] In addition, we create a dummy incentive for all positive R&D tax-credits since high-tech firms perform several activities which on average are affected by general incentives. We adopt this strategy because infrastructure is a common consumption good and we have a universe of public firms. In addition, we are also interested on firms' user cost. To this end, we rely on cross-state R&D data (see, for instance, Moretti and Wilson (2014) and Serrato and Zidar, 2018). Nonetheless, instead of using the expenditure share(s) from IRS which is equal to 0.5, we employ a longer time-span (1965-2015) from COMPUSTAT which is

[23]We have also collected other development incentives like job-creation tax credits from the same database. However, empirical evidence that looked specifically at the impact of those incentives and for instance job-training programs, did not find a positive significant relationship (Bremmer and Kesselring, 1993). In addition, when using the total amount of incentives the findings of Tuszynski and Stansel (2018) suggest a negative and significant relation in regards innovation outcomes (i.e., patents per 100,000 resident in that state).

a half of the suggested by (Wilson, 2009) without altering their main assumptions. Hence, our expenditure share is about half i.e. 0.1243 for the upper bound (90th percentile) between the above mentioned period. In this way, we control for the influence of potential outliers in our sample. Moreover, we conduct the same empirical exercise of Moretti and Wilson (2014) by regressing in the first stage the R&D user cost on state and year fixed effects and in the second, the predicted residuals on state R&D tax credit rate. Comparing the results in terms of the R-squared from that latter regression, our was 0.369 whilst theirs 0.345 which suggests that all within-state variability of monetary incentives comes from the interaction between R&D tax credits and firms' user cost.

Last but not least, we enrich our database with several labor market indicators from the BLS like civilian non-institutional state employment population, unemployment rate and population density using the Census maps. Also, we have taken the usual controls from the urban growth literature like average temperatures in winter (January), summer (July), high school and college graduation rates from total state population (Frank, 2009). Furthermore, we have also considered the effect of minimum wages changes over high-tech sectors. However, in the U.S for many states the minimum wage legislation allows that state levels go beyond than federal ones. As a result, we include a dummy variable equal to one for years in which state minimum wages were higher than federal ones. Data for historical state minimum wages comes from (Vaghul and Zipperer, 2016). Unlike earlier studies that focus on the effect of minimum wages outside tech sectors, here our scope is to control for minimum wage changes when both types of incentives (monetary and non-monetary) are simultaneously considered. In this way, we pick-up public policies effects over real wages for our two high-tech well defined groups (results are available upon request).

Finally, we take the Economic Freedom of North America (EFNA) subnational summary scores to account for the institutional environment among states (Stansel and McMahon, 2018, Tuszynski and Stansel, 2018). The index is composed of 10 variables divided into three areas: government spending, taxation, and labor market freedom and each category has its own indicator (e.g., labor market regulation score is baed on the Minimum Wage Legislation, Government Employment as a percentage of Total State/Provincial Employment and Union Density). At last, we control for supply side shocks by including a dummy variable equal to one when the economy had downturns periods (e.g., 1990-91; 2001 and 2008-2010) and zero in contrary case.

3.2 Descriptive Statistics and Correlations

In light of the empirical evidence of "superstar firms" and the U.S business slowdown, policymakers one may wonder to what extent infrastructure and State incentives (monetary and non-monetary) could have led to an uneven distribution in the production of knowledge (i.e., patents) in terms of aggregate value share. Furthermore, if between gap between top and bottom value-added high-tech firms implies a different response in employment growth and real wages as a result of the implementation of such public policies. Accordingly, prior to presenting the regression analysis, we will offer some descriptive discussion of the variables of interest in our database.

According to Table 1 public infrastructure (measured by per capita highway expenditures) has remained steadily with an average of 269 usd while private infrastructure (given by non-residential spending) accounting one-third approximately during 1993-2015 period. However, when comparing year-to-year public investments tend to be absorbed disproportionally more in bigger states (see, for example Figure A.1 in the Appendix). For what concerns employment rates and real wages in both top and bottom value-added high-tech sectors, results are diametrically opposed. In spite of a higher response of employment and wages —other things equal— in top value-added sectors with respect to bottom value ones; firms with a smaller contribution in terms of value share seem to provide more employment opportunities (i.e., longer intervals) and better salaries relative to more innovative sectors (see, for instance Figures A.2 and A.3 in the Appendix). Therefore, longer unemployment rates couple with lower wage growth rates due a fall in labor market fluidity (Davis and Haltiwanger, 2014). Another plausible explanation, as we will show later in our wage decomposition is a higher (though) heterogeneous response of conditional wages from high-tech workers in bottom value-added sectors. Furthermore, one of the novelties of this study is the inclusion of non-monetary incentives. From Table 1, our reduced policy average score is 0.047 with a maximum of 0.175 whereas our modified average user cost 1.144 which is not that far from the existing literature (Moretti and Wilson, 2014).

For what concerns innovation, Figures 2 and 3 show an uneven distribution of the aggregate value share as well as the production of innovation outcomes within high-tech sectors. These trends are driven by computer & electronic product manufacturing, publishing industries, telecommunications and professional scientific and technical services accounting more than 80% of patents with respect to other sectors and the same time the former delivering higher returns to the economy in terms of aggregate value. In addition, Figures 4 and 5 show that the average patents per 100,000 population at the 90^{th} and 75^{th} percentiles were 14 and 8 respectively whilst only 2.8 and 2.9 for the lowest percentiles 10^{th} and 25^{th} respectively. Interestingly, in Figures 6 and 7, a positive relationship between market power and the innovation gap is observed.[24] Simply put, the higher the knowledge gap (i.e., the difference between top and bottom percentiles), the lower will be the market competition because most innovative firms are more likely to absorb additional resources from generous R&D tax credits which ultimately provided them additional user cost saving effects as we will show below. Or else, as concentration rise, the aggregate value share reported by high-tech industries diminishes. We offer three possible explanations: i) knowledge is more concentrated in people; ii) innovation diffusion decreases with effort but the marginal return of new patents is more profitable in sectors with lower marginal costs and value share and iii) there is a higher dispersion on productivity (see, for example Akcigit and Ates, 2019*a*,*b*). Overall, this fact implies that the gap between these two groups has been widening-up especially after 1995 (see Figure A.4 and Figure A.5 in the Appendix).

To conclude, Table 2 shows the correlations between infrastructure, non-monetary incentives and conditional percentiles of innovation outcomes (i.e. patents and scientific citations) for each U.S census division. In panel A, a puzzling relation between infrastructure (public and private) and employment is observed. In the first case, highway expenditures seem to have a negative (but insignificant) impact in almost all divisions for top value-added industries but not for bottom ones (three out of nine were highly negative and significant). Conversely, there is a positive and significant correlation between non-residential spending and employment growth in New England, South Atlantic, Mountain and Pacific divisions. In regards to the diffusion of past public policies in several fields (third column), correlations are mixed but mostly positive rather than negative which could indicate that delayed effects over labor markets are more likely to occur. Conversely, in Panel B there is a positive and significant correlation between wages and non-monetary incentives in almost all divisions and again mixed correlations for both types of infrastructure expenditures. Although for the entire economy correlations are mostly positive for that reason we hypothesize ambiguous effects over high-tech wages. Finally, panels C and D show the correlations between wages and the (conditional) percentiles of technological patents

[24] As firms gain market power, labor share decreases; thus employment rates tend to be sluggish. This results in a higher dispersion on markups and a decrease on the output; thereby real wages will follow that trend because prices rise as well.

and scientific citations. To our surprise, both high-tech sectors seem to take advantage of public infrastructure investments for almost all percentiles while private infrastructure it not significant at all. In addition, the policy score is negatively correlated at the tails of innovation distribution but positive between the 25^{th} and 75^{th} percentiles. A possible explanation is that new ideas grow at a decreasing rate as innovation inputs have diminishing marginal returns (e.g., the value of creating new ideas is negatively related to time and scientists effort). Hence, regardless of their diffusion pace, the marginal contribution to labor markets of each new patent has less value in terms of aggregate knowledge. Certainly, these results do not imply that (future) benefits of new innovations will be passed-through to top value-added high-tech workers. The reason is that taxation effects among top innovators are heterogeneous. As a result, regardless the type of State incentives implemented by policymakers, highly qualified workers may inevitably abandon firms in the search of more profitable opportunities (the so-called "business stealing" effect). Nevertheless, we cannot attribute further explanations because correlation does not imply causation. For that we require a formal econometric model, which we address below.

Empirical Strategy

To account for the effects of infrastructure, incentives and innovation outcomes over regional labor markets for a specific group of industries (high-tech), this paper tests two well-defined and complementary hypothesis:

H_1 : *Infrastructure expenditures (public and private) as well as incentives: monetary (R&D tax-credits) and non-monetary (i.e., the mere diffusion of a policy) can be regarded as innovation boosters (e.g., a higher share of patents and scientific citations) for regional labor markets in terms of higher employment rates.*

H_2 : *The innovation gap (i.e., the difference between top and bottom percentiles) displayed by high-tech firms in terms of aggregate value share may be the result of an heterogeneous impact into the entire (conditional) wage distribution as public policies like infrastructure and innovation activities, depend on several and possibly opposed channels within each region (e.g., the proportion of high-school and college graduates, state minimum wage changes with respect to federal ones, etc).*

3.3 Identification Strategy

Given that highway capital outlays (X) are endogenous, it is most likely that errors can be correlated to our dependent variable. As a result, an exogenous instrument is required. Basically, the idea of an IV strategy is appealing: i.e., to use a variation in a third variable, Z (the instrument) that is exogenous (uncorrelated) with respect to the confounding variable, but at the same time correlated with X. Nonetheless, before considering any variable as a suitable candidate, two necessary conditions must be fulfilled: Relevance and Exogeneity. Put differently, the instrument needs to be relevant and foremost, it must be exogenous meaning that it affects the outcome variable only through the instrumented variable. Alternatively, the dependent variable is weakly affected by the potential instrument(s).

In empirical grounds, the relevance is usually assessed statistically thorough the Kleibergen-Paap F-stat statistic of first stage results. However, the endogenetity analysis must be complemented with other standard tests like the Hansen J-test (overidentifying restrictions), Anderson & Rubin (AR) statistic and the C test (orthogonality condition of instruments). As usual, rejection of the first test indicates the presence of weak instruments: the Kleibergen-Paap F-stat rule of thumb (for one endogenous regressor) suggests that its value should be above the threshold of 10. Conversely, the AR test can be decomposed into the K statistic and J-statistic, where the K statistic test if the exogeneity conditions are satisfied and the J statistic. However, the J statistics is evaluated at the null hypothesis that all instruments are exogenous as opposed to the J statistic from a GMM estimation, which is evaluated at the parameter estimate. Hence, the AR test provides more information about the exogeneity of instruments. At last, the orthogonality condition given by the C test shows that a subset of selected instruments are strictly exogenous. If rejection takes place, then, instruments may be weak.

In our case, we follow the path taken by Aghion et al. (2009), Aghion et al. (2019) and we will regard as an exogenous variation in infrastructure spending the number of senators and representatives in the appropriation committee (Stewart and Nelson, 2005, Stewart and Woon, 2017). As those studies claim, the number of committee members in the Senate is a powerful instrument for education expenditures and innovation. Even though the goal of above mentioned papers were the relation between education-growth and innovation-inequality, there is substantial evidence that it can be a useful (exogenous) and powerful instrument in regards to other state shocks like infrastructure expenditures (see, for example, Cohen et al. (2011) and

Hooper et al., 2018).

In fact, unlike more specific committees like transportation, appropriation committees not only have more explanatory power (its size is more than double with respect to transportation), but above all, only the appropriation committee of senators can authorize or grant additional funds to previously approved projects. As a result, in order to provide a more flexible approach, we recode the number of senators in the appropriation committee using the following classification: 0 if the State has no representatives, 1 with at least one representative, 2 if the State has two members, 3 if the State has at least three or more representatives.

Conversely, a different road-map would be to consider a projection-growth or a shift-share "Bartik instrument" reflecting the federal origin of spending funds excluding the region under consideration (Bartik, 1991). In plain words, an interaction between a time-varying shock (the shift: in this case real highway spending) and a cross-sectional sensitivity to the shock (the share of spending outside the state under consideration). In that vein, a new strand literature explores the potential benefits and drawbacks of this approach (see, for example, Goldsmith-Pinkham et al. (2018) and Borusyak et al., 2018).

Evidently, a large body of empirical papers still regard the "shift-share" approach as a valid strategy in the presence of endogenous regressors. Nonetheless, it may not be sufficient to ameliorate the endogeneity issue and it could even lead to the spurious regression problem. Indeed, interacting a time-series variable with another variable that varies only in the cross section provides flexibility but this comes at a cost, especially in the presence of common time-trends and/or shocks. What is more, placing higher weights on a particular sectors (e.g., nontradable vs tradable) may worsen both relevance and exacerbate the potential endogeneity (Broxterman and Larson, 2020) through a higher correlation structure between the errors and the shift-share instrument.

Accordingly, our empirical estimates relies on the committee members in the Senate as an exogenous variation of highway expenditures. As for the lag structure of the exogenous regressors, we follow the strategy applied by earlier studies Aghion et al. (2009), Aghion et al. (2019) and use a maximum of two lags. Yet, given that it takes time to a new member to affect highway outlays, we augment the number of lags of the endogenous variable up to 5. The latter is consistent with the fact that highways require maintenance, upgrading and completion based on pre-approved projects for which the senator was appointed.

4 Infrastructure, Incentives and Innovation

4.1 Employment Analysis

4.1.1 IV Results

H_1 : *Infrastructure expenditures (public and private) as well as incentives: monetary (R&D tax-credits) and non-monetary (i.e., the mere diffusion of a policy) can be regarded as innovation boosters (e.g., a higher share of patents and scientific citations) for regional labor markets in terms of higher employment rates.*

Before discussing the empirical results, we previously checked the stationary properties of all the variables involved in the analysis (results are available upon request). Additionally, in order to avoid the well-known spurious regression problem, we include the net GDP growth rate excluding both infrastructure measures. At last, a dummy variable equal to one for periods in which public capital outlays changed in terms of state population size is created. In this way, the stability of the panel is not compromised when Census division fixed effects are considered. It is worth stressing that we do not control for amenities like state crime rates because they are likely to be endogenous to current labor market conditions.

Tables 3 and 4, present the IV/GMM results using the above mentioned instruments for highway capital outlays. Throughout the employment analysis, I estimate models with infrastructure expenditures (public and private), incentives (monetary and non-monetary) and innovation outcomes measured by the disaggregated distribution of conditional patents in the high-tech industry with different aggregate value shares.

Columns I and II (both tables) with and without geography controls indicate the presence of a "market enabler" effect of private infrastructure over public infrastructure. That is in the short-run increase in private non-residential spending has a positive impact over employment. Point estimates in all high-tech industries and top value-added sectors are highly significant with an average value of 0.808% and 1.064% respectively. Conversely, in bottom value-added sectors results show a "disabler" and significant net effect. That is all coefficients of public infrastructure (i.e. with and without conditional patents) are negative and statistically significant with an average point estimate of -2.282% and -2.734% respectively whilst private infrastructure remains positive but statistically insignificant. Likewise, a policy diffusion score of past public policies has a positive and significant effect over total high-tech employment and top value-added sectors. For instance, a 1% increase in the policy score leads to an increase of 0.895% in employment for the top value added firms and 0.804 for the total industry other things equal. When innovation outcomes are considered, an inverted and significant u-shape is observed for all high-tech industries. Interestingly, these elasticities are driven by lower value-added sectors rather than top ones which remain insignificant with and without geography controls. As existing evidence on the U.S business slowdown documents, a possible explanation of these results can be the decreasing and/or stagnant aggregate value shares in sectors with higher marginal costs or more prone to experience higher labor turnover rates. In this way, the return of new ideas in sectors with lower value-added shares outperforms those with more aggregate value for the economy.

Moving further to column III, the "market enabler" effect with and without innovation outcomes vanishes as additional controls are included. That is point estimates continue to be positive and significant but the net effect of public infrastructure is stronger. Thus, the "disabler" net effect offsets the private market price effect regardless the group under consideration. Policy diffusion coefficients are slightly lower but continue to be positive and significant for the aggregate and top value-added high-tech sectors with the exception of bottom-value added which remain positive but statistically insignificant. For what concerns monetary incentives, more generous R&D tax credits offered by states, imply a positive and significant response on high-tech employment of 27.736% and 61.375% in top and bottom value-added sectors respectively for a three year (cumulative) window. Likewise, when conditional innovations are included the same trend is observed although point estimates in top value-added sectors cease to be significant. Along the same monetary lines, the R&D user cost shows that a 1% reduction is associated to a (15.769%) 21.681% employment increase in bottom value-added industries when innovation outcomes are (excluded) included, which is more than twice the amount of cost saving effects in top value-added sectors (-7.155%) and -10.935% respectively. When interactions between R&D tax-credits and firms user cost are considered, the difference between bottom and top value-added sectors remains significantly higher. That is point estimates of the interaction variable in high-tech sectors with lower value share are associated to a 17.889% of employment growth while in top sectors is less than half (8.166%).

Accordingly, our evidence shows that larger amounts of knowledge in most-innovative sectors do not necessarily imply a higher contribution to employment rates for the aggregate economy.

Expressed differently, the marginal return of a new patent and/or a scientific citation in the bottom value-added group has more value than the same innovation performed in the top value-added sector. Overall, the statistical performance of our instruments is acceptable given that both the overidentification and orthogonality tests (i.e., Hansen J-test and C-statistic) are not rejected at any significance level. This is also confirmed by the joint test of the structural parameters and the exogeneity of the instruments (AR test). In this case, the J statistic is evaluated at the null hypothesis as opposed to the Hansen J statistic from the GMM estimation, which is evaluated at the parameter estimate. Here rejection (at 10% level) takes place only in models without full controls and 5% to 10% when conditional patents are considered with full controls. In the remainder cases, the null hypothesis of weak instruments is above the 10% rejection interval. Nonetheless, caution must be taken with the latter results as they could be compromised by weak and strong cross-correlation an issue that we address below.

4.2 Robustness Checks

4.2.1 Instrumental Variable With Arbitrary Clustering

An important issue within panel data models that needs special care is serial correlation. In this vein, estimates can either be positively (negatively) correlated to the unobserved component of the dependent variable and/or with state-specific shocks among units and time (i.e., multiclustering). Moreover, persistence is likely to occur even in the absence of endogeneity; thereby, an adequate treatment of these problems is mandatory.

In empirical grounds, modelling the error term within a spatial framework provides a feasible strategy to tackle both issues. However, assuming a functional form of the error term may come with costs as multiway clustering implicitly regard a regularity condition. Simply put, each entry of the spatial matrix must share at least one dimension of clustering. This stringent but sufficient condition, means that both observations (in space and time) should be correlated with the error term and thus; they must also be dependent at time s. Unfortunately, in many real-life settings, this particular clustering structure does not hold.

Accordingly, an arbitrary cluster setting offers a more flexible approach since it allows units to be correlated with each other in any possible way, without any kind of imposed structure. What is more, regardless the data structure, not only it accounts cross section dependence and time dependence, but also interactions between the two dimensions as well. In this way, spatial changes can be captured over time or in any kind of decay between two moments in time t and s. For the sake of the space, technical details of the (adjusted) sandwich formula of the variance/covariance of the 2SLS estimator can be found in Colella et al. (2019).

In empirical grounds, an arbitrary correlation conveys that each observation's error term at a particular point in time may depend on other observation's error terms with a certain strength. All this information is collected in a matrix called (S). In the spatial context, S is normally built from information on the geographic distance between spatial units (e.g., latitude and longitude between regions, cities, and countries). It is worth stressing that unlike conventional inference —which leads to inconsistent estimates of the variance-covariance (VCV) matrix—, the performance rate in terms of rejection rates for small samples in arbitrary clustering frameworks is approximately 10% and converges quickly to the true significance level of 5% as the sample size increases. Therefore, this suggests that the latter correction produces consistent estimates of the VCV with respect to the former one. Thus, enabling applied econometricians to conduct robust inference in the presence of spatial correlation while accounting endogeneity.[25]

For what concerns our first research question, an instrumental arbitrary correlation framework is more suitable to account for different shocks like infrastructure expenditures, innovation outcomes and/or outliers. For instance, in the case of "star scientists" it is likely that contiguous states are affected by common shocks (e.g., tax-credits) and this should be reflected in the clustering structure. In the case of policy diffusion, clustering could be at work in first adopters and/or bordering states (e.g., tax and expenditure limits (spending limits): first year applied 1976; last year applied 1994). However, given that each state has its own tax legislation, and the results of such policies are publicly available, correlation among units will tend to decay over time. For the same token, the IV-AC is also robust to cases where public investments tend to be absorbed disproportionally more (i.e., bigger states) and at the same time, in states where there is an uneven number of representatives (i.e., senators) in the appropriation committee or innovation clusters.

[25]In the case of a shift-share analysis where the residuals between observations with similar distributions (of the shares) are correlated, this strategy would also be valid. Indeed, one would have to compute a similitude index between each pair of states based on a set of weights of the shift-share instrument(s) which yields a distance matrix S; and after that one could use S for arbitrary clustering. This refinement may alleviate a common problem of the shift-share empirical literature which is the relevance condition of an instrument with a higher weight.

4.2.2 IV Arbitrary Clustering Results

Tables 5 and 6 show the results of equation 3 accounting for an arbitrary clustering spatial framework. Columns I and II (both specifications) with and without geographical controls confirm that the "market enabler" effect of private infrastructure over public infrastructure still holds in all high-tech and top value-added sectors. That is point estimates are positive and highly significant. Likewise, in bottom value-added firms results confirm that the "disabler" effect of public infrastructure is still operating as coefficients of highway spending are negative and statistically significant. A policy diffusion score of past public policies has a positive and significant effect over total high-tech employment and top value-added but insignificant in sectors with lower value share. Interestingly, the upper part of the conditional distribution of innovations (i.e. patents) in top value-added sectors are statistically significant while in bottom value-added sectors the inverted u-shape effect weakness.

Accordingly, thus far the use an arbitrary clustering correlation with and without geographical controls delivers results which are quantitatively similar with respect to those in the previous section. Nonetheless, when more controls are included (column III), the "market enabler" effect without innovation outcomes vanishes (i.e. coefficients are no longer significant) and point estimates of public infrastructure are considerably lower in magnitudes and insignificant (-0.107 and -0.285) in all high-tech and top value-added sectors respectively. Conversely, when conditional percentiles of technological patents are considered private infrastructure is highly significant and offsets the negative effect of highway spending and the inverted u-shape relationship previously identified remains stronger for firms belonging to lower value share group rather than top value-added ones.

We offer two possible explanations for these new results: 1) the presence of "direct" networking effects of public policies. For example, a 1% increase in the the policy score implies a 0.793% employment growth for firms in top value-added sectors. Yet, the same argument does not hold for bottom value-added sectors given that point estimates continue to be positive but mostly insignificant. 2) increasing returns of monetary incentives and cost-saving coupled with "indirect" effects of State of past policies. For instance, entries in column III suggests that more generous R&D tax credits offered by states, imply a positive and remarkably different response on employment growth of 54.448% (85.034%) and 28.841% (27.184%) in bottom and top value-added sectors with and without innovation outcomes. Likewise, R&D user cost as well as interactions with generous tax-credit remain statistically significant and have the expected signs. Evidently, spatial interaction effects cannot be disregarded as these may explain why public policies like infrastructure expenditures and innovation policies are more effective in firms with different aggregate value shares.

In light of these new evidence some points are worth mentioning. First, the inclusion of incentives (monetary and non-monetary) coupled with infrastructure expenditures (public and private) and innovation outcomes (i.e. technological patents) must be considered "boosters" for other regional economies. In other words, cross-sectional dependence (CSD) is expected to be present in the analysis. For that reason, a weighted CSD test (CD_W) while accounting for serial correlation was employed (Juodis and Reese, 2021). Results show that the null hypothesis of (weak) CSD against strong one is rejected at 5% in GMM/HAC for all high-tech sectors whereas in the IV clustering framework the p-values of the CD_W are well-above the 10% rejection. Therefore, spatial effects and common factors play an important role in our analysis and these must be taken into account. Second, to have a qualitative response of how infrastructure, incentives and innovation outcomes impact high-tech workers, a wage decomposition is analysis necessary. We explore that venue in section 4.3.1. Third, interaction effects between spatial variables, incentives (monetary and non-monetary) and innovation outcomes may capture not only spillover effects but also agglomeration economies. To this end, below we apply filtering techniques and then check for the presence of CSD on the residuals of the Spatial Lag Model (SLX) while accounting for endogeneity.

Direct, Spillover Effects & Agglomeration Economies

In order to estimate direct and indirect effects we first test for the presence of spatial effects (Moran et al., 1993). Table 7 shows the results of Moran I spatial autocorrelation test for a 5 years window. As discussed earlier, we consider a window of five years because infrastructure investments are mostly slack and it takes time to unpack its cumulative effects. After confirming the presence of "local" spatial relations we estimate a SLX model. These type of models can be considered as a point of departure due to its flexibility in regards to direct and local spillover effects (see, for example Gibbons and Overman (2012), LeSage and Pace (2009), Halleck Vega and Elhorst (2015) and Elhorst (2021) among others).

Table 8 presents the results using the first-difference Bai-Ng filtering method. At first sight, after accounting for common effects, we obtain a lower CD statistic. That is p-values are well-above the 10% and with only one specification close to 9% rejection rate. Hence, we cannot reject the null of weak CSD hypothesis against strong one. Thus, estimates do not suffer from spurious inference problem. Furthermore, our SLX model includes interactions between exogenous regressors and spatial lag variables. These new variables capture not only indirect effects and/or agglomeration economies, but also account the spatial endogeneity. It is worth mentioning that we do not instrument spatial lagged variables. The reason is spillovers effects generate an obvious violation of the exclusion restriction. That is if other predictor(s) values directly affect the outcome they cannot be valid instruments. In the case of a public goods like infrastructure the inclusion of past predictors of the spatial variable produce a tension because the other spatial relationships must be rule out simultaneously. That is units should not be benefited from these investments which is an unrealistic assumption. Moreover, if the spatial instrument is exogenous it would imply that there are no simultaneous relationships (between units) involved in the analysis both in the first and second stage analysis. Accordingly, an instrumentation of the spatial lag variable should be avoided as it constitutes a violation to the system of equations given that variation comes from within and not from the outside.

The SLX results show that direct effects of public infrastructure are positive and significant at 10% level. That is coefficients rang from 0.129 to 0.154 indicating that after 8 years highway expenditures increase high-tech employment. Surprisingly, private infrastructure measured by non-residential spending is insignificant or most probably positive effects are located for more disaggregated units (e.g., counties). In regards to innovation, both sectors top and bottom value-added show a positive significant response. Particularly, a priori one would expect that firms with a higher aggregate value share would contribute more to high-tech employment than those with lower value. However, when spatial factors are accounted for point estimates are both positive and significant but remarkably close each other: 0.616 and 0.510 respectively. This fact confirms my previous findings that is the return of new ideas is more profitable in sectors where marginal costs are already lower.

Last but not least, I find no evidence of public infrastructure spillovers when local revenues for roads, streets, and highways increase. This fact can be explained by a higher presence of representatives in the appropriation committee. Expressed differently, one would expect that states with more senators may influence the budget to increase the performance not only on high-tech industries but also in other traditional industries (like construction) as well. But, as previous studies claim, in the short-run infrastructure effects are slightly positive and/or close to zero e.g., for each dollar spent on the construction sector in the year 2010, only six jobs were created per $1 million (Garin, 2019).

Interestingly, interactions between non-monetary incentives (i.e. the diffusion of past policies) and innovation outcomes in both sectors are both positive (0.183 and 0.152 respectively) but only significant in sectors whose aggregate value share is larger. That is networking effects of ideas are transferred to other states provided that firms stimulate their local economies by increasing rates of ideas (i.e. production of technological patents) due to generous tax-credits. This is an important contribution of my paper because it links the effectiveness of past public policies, innovation outcomes and labor market performance in sectors where agglomeration economies and externalities are considerably larger with respect to construction industries usually associated to these investments.

Finally, I test for the presence of indirect effects of tax-credits through innovations (i.e. patents) in both sectors and firms' user cost. Compared to the existing evidence (see, for example Moretti and Wilson, 2014), I find positive and significant spillover effects over total high-tech employment. For instance, when interactions between R&D user costs technological patents in bottom value-added sectors are considered, the effect of generous subsidies is 0.431 and significant while in sectors with higher aggregate value share the effect is slightly lower (0.424) and also statistically significant. This is an important departure from the "beggar the neighbour effect" (see, for instance Wilson, 2009) because infrastructure investments couple with incentives both (monetary and non-monetary) indirectly generate positive spillover effects in other high-tech industries and foremost curbs inventors incentives to move out from their regions. For those reasons, the evidence on this paper supports the effectiveness of "jobs to people" policies in regards to business dynamism for regional economies.

4.3 Wage Analysis

4.3.1 Quantile Wage Decomposition

Extant literature of road infrastructure has applied quantile methods in the past but with a focus on poverty (Khandker and Koolwal, 2010, 2011). To be precise, the aforementioned studies employ a correlated random component as an additional covariate in their models to account for unobserved heterogeneity amongst units. However, a problem with this approach is that including control variables in a quantile regression model alters its interpretation for a specific individual with and without a predetermined characteristic.

To empirically address our second question, the within-transformation could be used to wipe out all the effects across variables before applying OLS. Unfortunately, this strategy fails to uncover unobserved heterogeneity and above all, its bias will be related to the unobserved component of model. That is the τ^{th} quantile for a particular state with lets say better infrastructure services and/or with more innovation clusters is not the same as the τ^{th} quantile for another one with lower quality and/or lower quantity of patents per state. Another alternative is to simply take the pooled averages of the population i.e., averaging the individual-level effect across units. But, in this case the average (marginal) effect will be constant as we assume that the model is linear on its parameters; thereby, we will not be able to capture the unobserved heterogeneity.

Accordingly, from an empirical point of view there is substantial evidence that the use of quantile regression models overcomes both problems and above all it is robust to the presence of outliers. Nonetheless, as in any scientific subject, not every problem can be addressed with the same recipe. In other words, there are several quantile estimators which make use of different assumptions to recover the parameters of interest and at the same time exploit the conditional heterogeneity among units.

Following the previous line, the simplest approach would be a linear (additive fixed effect) estimator which appropriately takes care of the unobserved heterogeneity. In this line, as Canay (2011) suggests, the problem is sequential. In the first step, all the individual effects are estimated by using the standard (within) fixed effects estimator for linear panel data. Later in a second step, the recovered coefficients are estimated by standard quantile regressions (i.e., treating the estimated individual effects from the first stage as given). Overall, the intuition behind this procedure is straightforward because the researcher assumes that the parameters will remain unchanged.

Nonetheless, the procedure from above is not exempt from drawbacks. On the one side, as Chen and Huo (2019) points out, simplicity comes at costs. First, the two-step procedure requires certain moments of the idiosyncratic errors to exist, and thus the robustness of quantile regressions against heavy-tailed distributions must be sacrificed. Second, the assumption that $N/T^s \rightarrow 0$ for some $s > 1$ is not enough to ignore the asymptotic bias of the estimated coefficients. For example, according to their simulations, coverage rates of the confidence intervals (with 95% nominal levels) based on this type of estimator are lower than 1% when N = 1,000 and T = 20 which are not that far from our data structure (N=1,300 and T=26). In addition, the asymptotic variance of the Canay estimator for the constant term is not correct. To this end, the authors suggest a smooth quantile regression (SQR) instead of the quantile regression to estimate the coefficient of the regressors while preserving the first stage fixed-effect estimates. In this way, coverage rates of the confidence intervals are improved once the analytical variance correction is done; thereby providing a valid inference of the estimates.

On the other side, in the context of a continuous (non-discrete) endogenous regressor like highway capital outlays, if the dependent variable and/or its lags are correlated with the variable being instrumented, then, the IV moment condition considered in the approach of (Chernozhukov and Hansen, 2008) may not be appropriate. Likewise, under the presence of individual factors correlated with the independent variable(s), one could apply the estimator suggested by Harding and Lamarche (2009). Their quantile estimator allows the endogenous variable to be correlated with unobserved factors affecting the conditional response variable. In this way, the researcher can estimate the covariate effects at different points of the distribution while controlling for individual heterogeneity.

Certainly, the estimators from above have their own benefits and pitfalls because including a set of dummy variables representing each cross-sectional unit could create an incidental parameter problem affecting the consistent estimation of all coefficients in the model (further details are discussed in Koenker (2004), Powell (2016) among others). Moreover, since fixed effects are additive we cannot make interpretations at an individual level. Although, we might report the "average" conditional quantile effect across the whole population as representative effects assuming that the parameters of interest do not vary based on the fixed effects (in the spirit of Canay two-step estimator).

Consequently, for our wage decomposition model we proceed with the quantile panel estimator with nonadditive fixed effects proposed by Powell (2016). The intuition and simplicity of this estimator makes the computational analysis less burdensome. In plain words, it assumes that the parameters of interest vary based on the nonadditive disturbance term while maintaining the nonseparable disturbance property of the basic quantile regression model which is the backbone of quantile regression models. Yet, there are two major concerns that must be taken into account.

First, the existence of large number of fixed effects to be estimated and second; the incidental parameters problem when the dimension of the panel is relatively small. In broader terms, the estimator from above overcomes both issues to our data generating process; thereby each regressor can be recovered if there is enough variability within each iteration (below we will briefly outline its intuition).

To motivate further our empirical application for high-tech industries, we follow the notation of Powell (2016). From an empirical point of view, our wage decomposition aims to shed some light on the tech/non-tech puzzle found by Kemeny and Osman (2018). However, in our case we hinge upon the empirical regularities of U.S public firms for which different aggregate value shares and hypothesize that such puzzle may be the result of a set of past public policies like infrastructure expenditures and incentives (monetary and non-monetary) that could have affected the conditional distribution of wages in (within) high-tech sectors.

Formally, our model can be specified as follows

$$y_{it} = D_{it}\mu`(U_{it}^{*}) \quad U_{it}^{*}\, U(0,1) \tag{4}$$

where y_{it} is the outcome variable which is log annual payroll of high-tech industries (top and bottom value-added per 100,000 population); $D_{it} = (I_{it}, X_{it})$ is a vector of treatment variables for each state i at time t. In addition, we include a set dummy variables like Census divisions, recessions, state minimum wage changes with respect federal ones, time effects (e.g., decades to capture cumulated effects) as well as our monetary and non monetary incentives (i.e., infrastructure, the diffusion of policies and innovation outcomes) along with the usual controls from the urban economics literature. It is worth stressing that U_{it}^{*} is general disturbance term which may be a function of several disturbances. Since it is uniformly distributed, we can say that it represents "proness" for the outcome variable. Therefore, we can express it as an unknown function of individual fixed effect (αi) and the observation specific disturbance term U_{it} as follows

$$U_{it}^{*} = f(\alpha i, U_{it}^{*}) \quad where \quad U_{it} \sim U(0,1) \tag{5}$$

With respect to traditional quantile methods which assume independence between U^{*} and D, the nonadditive fixed effect estimator does not specify nor estimates the individual fixed effects. Accordingly, the number of parameters to be estimated is relatively small of other estimators like Canay (2011).

In practical terms, the implementation of the nonadditive estimator has two important advantages with respect to the quantile fixed-effect estimator. First, it provides the estimates of the conditional distribution of $yit|D$ which can be interpreted with the usual standard inference tools as in many empirical applications. Second, given that fixed effects quantile regression estimates are in general biased and inconsistent when the time dimension of the panel is small and fixed (the so-called incidental parameter problem), once identification is achieved, the nonadditive quantile panel estimates are consistent, have an asymptotic normality and foremost they provide an adequate inference (Powell, 2016). Furthermore, even with a short time span (e.g T=2) the estimator still has a good performance.

Finally, it is worth mentioning that direct comparisons between both quantile specifications (with and without additive fixed effects) should be made with caution because the latter is estimated through the Markov Chain Monte Carlo (MCMC) optimization method which is sensitive to the order of the regressors as well as the acceptance rate of the algorithm implemented. Also, the casual inference of the estimates with additive fixed effects is completely different.

Normality test for panel data models

Normality of the error structure term is one the main assumptions for conditional mean regressions. Therefore, it is meaningful to check whether or not the OLS model is suitable for the data generating process (DGP). Specifically, if the DGP is normally distributed, then, parametric linear mean regression models could be adopted to investigate the impact of infrastructure and innovation policies over high-tech wages. Conversely, if normality does not hold, then, OLS methods are not suitable as conditional means would provide biased and non-robust estimates. In this situation, panel quantile regression models are more appropriate as they can naturally deal with the unobserved heterogeneity of the DGP.

From an empirical point of view, there are two ways to check for the normality assumption. On the one hand, one could use a quantile-quantile plot (Q-Q plot) to have a quick glimpse of the DGP. Indeed, if the variable(s) are normally distributed, then, the scatter diagram(s) should be on the diagonal line. Rejection takes place when diagrams for each Q-Q plot deviate from the diagonal lines. Another interesting and (practical) approach is to visually assess the skewness and kurtosis of the variable(s) of interest. In our case, we adopt a two-step procedure and first visually evaluate normality by focusing on pairs of relations between wages, innovation and infrastructure. For the sake of the space, we refer the reader to Figures A.6, A.7, A.8 and A.9 in the Appendix. Overall, results from the analysis show fatter (skewed) tails by decades and divisions.

In a second step, we adopt a formal test to check whether or not quantile methods are suitable. In plain words, normality is tested separately and jointly through the innovation and specific error terms (for more details, see, Galvao et al., 2013). More precisely, the test controls for normality within the panel structure by assessing skewness and kurtosis. Overall, rejection of the null hypothesis would indicate that the error structure does not follow a normal distribution. Hence, the linear regression model (LR) based on the conditional mean estimation would not be able to account for the unobserved heterogeneity. At last, disregarding this valuable information in terms of deviations from (a)symmetry and/or excess of kurtosis may cast doubt on the empirical estimates.

We perform the analysis for our two high-tech sectors using the same controls from the employment models but including dummies by regions and decades. Results in Table 9 confirm that the null hypotheses of the 6 indices are rejected. Indeed, skewness is highly significant in both error terms (innovation and state) whereas individual kurtosis is significant at a 10% in bottom value-added sectors. Nonetheless, both joint hypothesis are rejected at 1% and 5% level respectively. Therefore, based on these results, we conclude that wage distributions in both groups (top value and bottom value-added) are asymmetric and left-skewed.

Overall, both strategies (a visual inspection and a formal test) imply that our sample data is not normally distributed. Consequently, OLS methods are not appropriate to investigate the effect of infrastructure and innovation policies over high-tech wages. Therefore, the panel quantile regression with nonadditive fixed effects is more appropriate for our wage decomposition. Notwithstanding the clear differences between conditional mean regressions and quantile panel data models, we will compare their performance to double-check our results. Nonetheless, caution must be taken because the latter is estimated through the Markov Chain Monte Carlo (MCMC) optimization method which is sensitive to the order of the regressors.

Interestingly, another striking difference of the above mentioned framework with respect to standard quantile models (i.e. with additive effects) is the casual inference of their estimates. In the second case, the interpretation is completely different because it separates the disturbance term into different components and assumes that the parameters vary based solely on the time-varying components of the error term. Put differently, it alters the interpretation of the parameters of interest relative to cross-sectional quantile regression (QR). Hence, if non-separability holds, then, the resulting estimates can be interpreted in the same spirit as cross-sectional quantile estimates (i.e., what would be the impact on the quantile τ^{th} due to a change on the explanatory variable). For all those reasons, our empirical estimation implements the instrumental quantile estimator with nonadditive fixed effects.

4.3.2 Quantile Wage Decomposition With Nonadditive Fixed Effects Results

H_2 : *The innovation gap (i.e., the difference between top and bottom percentiles) displayed by high-tech firms in terms of aggregate value share may be the result of an heterogeneous impact into the entire (conditional) wage distribution as public policies like infrastructure and innovation activities, depend on several and possibly opposed channels within each region (e.g., the proportion of high-school and college graduates, state minimum wage changes with respect to federal ones, etc).*

In lieu with the empirical wage literature, we measure the dependent variable in log values and normalized its value by state population to control for business cycle effects. As a first step, we examine the performance of the quantile estimator with nonadditive effects by running a simple model without controls. More precisely, we include both measures of infrastructure (public and private), non-monetary incentives (i.e. a policy score) and conditional patents controlling only for dummies (U.S Census division and decades). As explained, these effects will be separated from the main covariates, providing us a clean inference. Again, we check for normality applying the above mentioned test. Results (available upon request) show that the null hypothesis of 6 indices for top value-added sectors and 5 in bottom value-added ones are rejected. Hence, wages from high-tech workers are asymmetric and left-skewed.

In Figure 8 we provide the marginal effects of our benchmark specification for the most important quantiles. At first sight, the "disabler effect" of highway spending is also at work regardless the group under consideration. Indeed, in top and bottom value-added groups the negative impact decreases throughout quantiles. Not surprisingly, the market price effect of non-residential private expenditures is positive and highly significant at the 50^{th} quantile and beyond the median for both sectors. In plain words, private infrastructure ameliorates the negative initial impact of public investments of workers located in the median and above the 75^{th} and 90^{th} quantiles in bottom value-added sectors. For what concerns innovation, results are mixed: A higher number of high-tech patents is associated to better salaries in both sectors throughout quantiles. In line with empirical evidence (see, for example, Moretti and Wilson, 2014), top earners seldom enjoy the full benefits of their own scientific contributions. For instance, a 1% increase in technological patents workers leas to a 0.075% and 0.096% in wages for both top and bottom value-added sectors respectively. Lastly, a higher policy score is positively associated o higher wages in top value-added sectors while in bottom value-added the relationship is mostly negative and significant with the exception of top earners. Nonetheless, as we will show later, these results do not prevent that workers from those industries may enjoy from better salaries due to monetary incentives provided by states (e.g., R&D tax incentives) and/or higher minimum wages with respect to federal ones. Accordingly, with respect to conditional mean estimates (not reported but available upon request), quantile models unpack significant (though heterogeneous) effects of infrastructure, non-monetary incentives and innovation outcomes across the entire (conditional) wage distribution of high-tech industries.

Moving further to a full specification, we estimate a quantile panel data model with nonadditive fixed effects using the same controls employed in the employment analysis. Before discussing these results, we first check for the absence of normality for the two groups of high-tech industries. In Table 9 the null hypotheses of the 6 indices for top value-added sectors are rejected while 5 indices in the bottom value-added ones. Again, skewness is highly significant in both error terms (innovation and state) whereas state specific kurtosis ceased to be significant in the latter group. Nonetheless, both joint hypothesis are rejected at 1% and 5% level respectively; thereby indicating a leptokurtic distribution. Therefore, we conclude that distributions from both groups are asymmetric and left-skewed.

Table 10 shows our main results of the wage decomposition. The negative and significant impact of public infrastructure is wiped-out after the 20^{th} quantile in top value-added sectors. Interestingly, workers located at the top side of the wage distribution (95^{th} percentile) are more likely to enjoy the benefits of highway expenditures. Evidently, quantile models are more flexible to unpack the effectiveness of these policies at different parts of the distribution with respect to conditional (mean) estimates where coefficients of public infrastructure are positive (0.025%) but statistically insignificant. Conversely, in bottom value-added sectors a sort of inverted u-shape relation is observed. For instance, a 1% increase in highway spendings is associated to a 0.189% increase in salaries for workers located in the upper side of the distribution (95^{th} quantile). Again, conditional mean models indicate a decrease though insignificant effect of 0.120% in wages. In the case of private infrastructure, for top value-added sectors there is a positive effect of non-residential spending throughout quantiles with a maximum of 0.127% at the 75^{th} quantile. Oppositely, in bottom value-added sectors private infrastructure is mostly negative and statistically significant specially at the 10^{th} quantile where point estimates indicate a decrease of -0.255% in real wages. However, as we move across the distribution, its magnitude weakens to finally become positive but insignificant i.e. 0.197% at the 95^{th} quantile.

For what concerns innovation, results results are mixed. On the one side, in top value-added sectors a higher number of technological patents and scientific citations is positively associated to better salaries. However, in the upper side of the distribution coefficients become negative and cease to be significant. Conversely, in the bottom value-added group the relationship is mostly monotonic and highly significant. Again, as in the benchmark case, a possible explanation is the presence of decreasing marginal returns in regards to new knowledge and/or displacement effects amongst top earners in sectors whose aggregate value share is higher. Conversely, in bottom value-added sectors these returns are more profitable because each new patent makes a substantial improvement to the industry in terms aggregate value share. Expressed differently, the weights of patents and wages are inversely related in top value added sectors. For those reasons, we argue that top innovators are more likely to abandon firms in the search of better opportunities (the so-called "business stealing" effect).

In regards to non-monetary incentives, a higher policy score does not necessarily imply better salaries. In top value-added sectors there is positive effect over wages but these are restricted only to the tails: 0.029% and 0.005% in the 10^{th} and 20^{th} quantiles respectively. Nonetheless, throughout quantiles point estimates start decreasing from the median onwards to finally become negative and statistically significant. Oppositely, in bottom value-added industries coefficients are mostly negative and significant suggesting that in the medium-run, the diffusion of past policies are associated to lower salaries. In that vein, two points are worth stressing. First, the fact that innovation networks may have had positive effects over employment rates does not imply that workers from those industries would have been benefited proportionally from such policies. Second, the universe of high-tech industries is composed by heterogeneous sectors with different marginal costs and technologies. Hence, it seems reasonable to expect that wages may also depend on specific job characteristics (e.g., creativity, originality, etc) as well as schooling requirements (i.e., high-school or college degree) involved in tradable and nontradable sectors like in Kemeny and Osman (2018) framework. Evidently, a comprehensive study of infrastructure expenditures considering all those aspects comes at the cost of more disaggregated relations between firms and workers. Hence, we leave that for future research.

Last but not least, we find important effects of general R&D tax-credits on wages but magnitudes cannot be compared directly to those from the existing literature (Moretti and Wilson, 2014). Naturally, one may wonder whether or not the employment benefits found in our empirical analysis were also pass-through to better salaries in industries with different aggregate value shares. In the case of top value-added sectors, our quantile results show a significant effect of 2.005% at the tails (10^{th} quantile). However, as we move across the distribution, point estimates become negative and significant e.g., -4.432% at the 95^{th} quantile (top earners). This result reinforces the fact that most qualified workers seldom enjoy the full benefits of development incentives. For bottom value-added sectors, these generous subsidies are mostly associated with higher salaries in almost all quantiles, especially at the 50^{th} quantile (4.572%). Therefore, the median worker benefits more from R&D tax-credits and has on average a higher salary with respect to another same worker in top value-added sectors. Accordingly, monetary incentives are positively (negatively) associated to higher (lower) salaries in bottom (top) value-added industries which suggests that these generous subsidies do not always lead to higher wages. Likewise, both R&D user cost and interactions display mixed results. That is the direction of coefficients might lead to payroll reductions or increments depending on the quantile and/or sector under analysis. For example, at the tails of the wage distribution (10^{th} quantile), a 1% reduction in the R&D user cost results in a 0.609% and 2.204% increase in the average salary in top and bottom value-added sectors respectively. However, for top earners these are considerably much larger in bottom (14.779%) rather than top value-added (0.148%) sectors. Interestingly, interactions between R&D user cost and state subsidies suggest that only workers located at the 20^{th} quantile benefit from a 1.353% increase in real wages. Conversely, in top value-added sectors point estimates are mostly positive and statistically significant for top earners (e.g., 1.210%).

In light of these new results we offer two plausible explanations. First, development incentives have larger effects on firms whose marginal costs and/or aggregate value share are already lower. This would explain why top earners (i.e. those with more human capital and skills) from top value-added sectors are more likely to abandon firms in the search of more profitable opportunities and not the other way around. Unfortunately, one limitation of using aggregated data is that we cannot account the proportion of workers belonging to a top or bottom value-added sectors willing to leave the firm. Second, the presence of labor market regulations like minimum wages could be at work. To this end, we explore that venue and re-run our wage decomposition including a dummy variable equal to 1 for changes in state minimum wages with respect to federal ones. This is supported because each state is relatively an independent unit and may have a different wage legislation. Results (available upon request) show a positive and significant impact over real wages in almost all quantiles. In the case of top value-added sectors, higher minimum wages with respect to federal ones imply better salaries for workers located at the 10^{th} (0.117%) quantile while for top earners the effect is moderate (0.083%) but continues to be statistically significant. For bottom value-added sectors, coefficients are larger and significant but beyond the median the magnitudes drop; and for top earners, point estimates become negative and significant. Likewise, the majority of R&D tax-credit coefficients thorough quantiles in both high-tech groups are negative and statistically significant. These results could be framed into the existing empirical evidence of longer unemployment rates couple with lower wage growth rates due a fall in labor market fluidity (Davis and Haltiwanger, 2014).

All in all, results are consistent with a model in which the labor supply of high-tech workers at the state level is sensitive or elastic. The empirical approach applied in this paper unpacks sizeable heterogeneous effects of infrastructure expenditures, incentives (monetary and non-monetary) and innovation outcomes measured as the quantity of technological patents and scientific citations for high-tech industries with different aggregate value shares. Usually, such heterogeneity cannot be captured with conditional mean estimates because of the strong assumptions behind linear models. From an empirical point of view, these generous subsidies involve several and complex relations between firms and workers regardless the contribution of each sector in terms of aggregate value to the economy. Accordingly, in order to have a better understanding of how these public policies work at a firm level, more research is necessary.

5 Concluding Remarks

This paper examines the effects of regional employment and wages from infrastructure expenditures (public and private), incentives (monetary and non-monetary) and innovation outcomes (i.e. patents and scientific citations) for U.S high-tech industries. One of the novelties of the present paper is the inclusion of non-monetary incentives (proxied by the diffusion of past public policies applied since the 1913's) and the production of ideas (i.e. technological patents and scientific citations) conditional on industry subsectors value share. While in the past other studies focused on event study programs (e.g., the ARRA) or generous tax subsidies towards specific industries (e.g., biotech), this paper brings new perspective on the effectiveness of public policies for regional labor markets. In order to assess the effects of infrastructure, incentives and innovation outcomes, I first identify the sectors for which infrastructure and innovation policies have the largest effects in terms of aggregate value share and different forms of agglomeration economies and thick market externalities (new jobs and better salaries): high-tech industries like computer and electronic product manufacturing, telecommunications and ICT subsectors. Then, I propose a simple regional model in which workers have some kind of mobility and firms make use of both type of incentives: monetary (e.g., tax-credits) and non-monetary (the mere diffusion of past public policies) and infrastructure expenditures (public and private) to highlight the channels.

The first contribution of the paper is the identification of important stylized facts from the U.S business slowdown (e.g., higher knowledge gap and lower market competition between top and bottom innovators in public firms). That is, firms with lower marginal costs are more likely to be benefited from these public policies. The second contribution is an empirical assessment of the channels for which the above-mentioned policies work in regional labor markets. For employment models (under an IV arbitrary framework) and using the number of senators in the appropriation committee as instruments, I find that a 1% increase in the diffusion of past policies leads to a 0.777% increase in employment for the top value added sectors other things equal. In addition, a 1% increase in (public infrastructure) private non-residential spending is associated with a 0.941% increase (decrease of -2.034%) in employment for top value-added (bottom value-added) when conditional innovation outcomes are considered. That is an "enabler" entrepreneurship effect over total and top value-added high-tech employment due to private infrastructure expenditures and a "disabler" negative effect of highway spending only in sectors with lower value share.

Interestingly, the cumulative response of monetary incentives (both generous subsidies as well as firms' user cost) are positively associated to higher job creation rates in bottom rather than top value-added sectors. Contrary to traditional wisdom, R&D tax credits offered by states, imply a positive and (cumulative) response on employment of 20.841% and 54.448% in top and bottom value-added sectors respectively while a 1% reduction in the R&D user cost leads to a 22.703% increase in employment in bottom value added sectors. In the top value added group and industry as a whole, point estimates are statistically significant but less than half compared to those in lower value share sectors (9.225 and 11.553) when conditional patents are considered. Therefore, by assuming a flexible relation among units and time, results show that larger amounts of knowledge in most-innovative sectors do not necessarily imply a higher contribution to employment rates for the aggregate economy. When spatial variables are introduced into the analysis, I find positive but moderate direct effects of public infrastructure. That is coefficients rang from 0.129% to 0.154% indicating that after 8 years highway expenditures increase total high-tech employment. Moreover, including spatial interactions between non-monetary incentives and innovation outcomes in high-tech sectors with a higher aggregate value share leads a 0.183% employment growth i.e. networking (spillover) effects in other states provided that firms stimulate their local economies through increasing rates of ideas (i.e. production of technological patents) due to generous tax-credits.

Finally, a wage decomposition for within high-tech workers is performed. Results confirm the same trends previously observed in the employment analysis. Private infrastructure spending has a positive effect across subsectors and throughout quantiles while the negative impact of public infrastructure is wiped-out after the 20^{th} quantile in top value-added industries. For instance, a 1% increase of highway expenditures in the upper side of the wage distribution leads to a positive and significant impact of 0.041% and a 0.189% in top and bottom value-added sectors respectively. In addition, monetary incentives are positively associated to higher wages in top value-added for workers located at the tails of wage distribution (e.g., 10^{th} quantile) and highly significant. However, as we move across the distribution coefficients become negative and significant, especially for top earners (-4.432%). Conversely, high-tech workers from bottom value-added sectors have on average better salaries in almost all quantiles, especially at the median (4.572%) compared to workers in top value-added sectors. Accordingly, monetary incentives are positively (negatively) associated to higher (lower) salaries in bottom (top) value-added industries which suggests that these generous subsidies do not always lead to higher wages. In the case of both R&D user cost and interactions, results are mixed and should be taken with caution as the presence of labor market regulations like minimum wages may be at work. Therefore, regardless the contribution of each sector in terms of aggregate value to the economy, we retain that more research is necessary because high-tech industries are continuously evolving and the labor supply of high-tech workers is highly elastic.

From a state's perspective, efforts should be concentrated on the empowerment of laggard high-tech sectors and the reduction of the innovation/wage tech gap in terms of aggregate value. Certainly, this would require not only more job-training programs, but also investing these generous R&D subsidies and federal grants into new low-cost technologies and/or dynamic processes based on worker's capabilities with soft skills. From an institutional perspective, it is well-known that States spend billions of dollars to attract R&D activity to their jurisdiction(s) to enhance their local economies. The direction of these generous incentives have always been a subject among debate among policymakers because these can generate different forms of agglomeration economies and thick market externalities at both margins (intensive and extensive) for regional labor markets. To the best of my knowledge, this paper shows that infrastructure expenditures coupled with monetary incentives like R&D tax credits and non-monetary incentives like the diffusion of past public policies, may facilitate the transition to achieve such targets in high-tech industries with different aggregate value shares. Unlike earlier studies (Goolsbee, 1998), we sustain a more optimistic perspective of innovation policies in the sense that to have well-remunerated workers, substantial effort by firms (in terms of cost-saving effects) must be done to hire additional workers and provide better salaries. Despite finding networking effects in regards to innovation, the idea that policymakers can stimulate one specific industry (high-tech) and not another one is always questionable given workers' tastes and degree of mobility. For that reason, we hasten to point out whether or not should States support these generous tax-credits or else the minimum conditions in terms of the production of knowledge (patents) relative to the aggregate value share generated by firms to apply for these grants as more research is necessary to understand of how these public policies work at a firm level.

References

Acs, Z. J., Audretsch, D. B. and Lehmann, E. E. (2013), 'The knowledge spillover theory of entrepreneurship', *Small Business Economics* **41**(4), 757–774.

Acs, Z. J., Desai, S. and Hessels, J. (2008), 'Entrepreneurship, economic development and institutions', *Small business economics* **31**(3), 219–234.

Aghion, P., Akcigit, U., Bergeaud, A., Blundell, R. and Hémous, D. (2019), 'Innovation and top income inequality', *The Review of Economic Studies* **86**(1), 1–45.

Aghion, P., Boustan, L., Hoxby, C. and Vandenbussche, J. (2009), 'The causal impact of education on economic growth: evidence from US', *Brookings papers on economic activity* **1**, 1–73.

Akcigit, U. and Ates, S. T. (2019*a*), Ten facts on declining business dynamism and lessons from endogenous growth theory, Technical report, National Bureau of Economic Research.

Akcigit, U. and Ates, S. T. (2019*b*), What Happened to U.S. Business Dynamism?, Technical report, National Bureau of Economic Research.

Akcigit, U., Grigsby, J., Nicholas, T. and Stantcheva, S. (2018), Taxation and innovation in the 20th century, Technical report, National Bureau of Economic Research.

Amior, M. and Manning, A. (2018), 'The persistence of local joblessness', *American Economic Review* **108**(7), 1942–70.

Armstrong, H. W. (1995), 'Convergence among regions of the european union, 1950–1990', *Papers in Regional Science* **74**(2), 143–152.

Arora, A., Belenzon, S. and Sheer, L. (2017), Back to basics: why do firms invest in research?, Technical report, National Bureau of Economic Research.

Audretsch, D. B., Heger, D. and Veith, T. (2015), 'Infrastructure and entrepreneurship', *Small Business Economics* **44**(2), 219–230.

Autor, D., Dorn, D., Katz, L. F., Patterson, C. and Reenen, J. V. (2017), 'Concentrating on the Fall of the Labor Share', *American Economic Review* **107**(5), 180–185.

Autor, D., Dorn, D., Katz, L. F., Patterson, C. and Reenen, J. V. (2020), 'The fall of the labor share and the rise of superstar firms', *Quarterly Journal of Economics* **135**(2).

Bai, J. (2004), 'Estimating cross-section common stochastic trends in nonstationary panel data', *Journal of Econometrics* **122**(1), 137–183.

Bai, J. and Ng, S. (2002), 'Determining the number of factors in approximate factor models', *Econometrica* **70**(1), 191–221.

Bai, J. and Ng, S. (2004), 'A panic attack on unit roots and cointegration', *Econometrica* **72**(4), 1127–1177.

Bailey, N., Holly, S. and Pesaran, M. H. (2016), 'A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence', *Journal of Applied Econometrics* **31**(1), 249–280.

Bartik, T. J. (1991), 'Who benefits from state and local economic development policies?'.

Bartik, T. J. (2017), 'A new panel database on business incentives for economic development offered by state and local governments in the United States'.

Bartik, T. J. (2019*a*), 'Should Place-Based Jobs Policies Be Used to Help Distressed Communities?', *Available at SSRN 3430374* .

Bartik, T. J. (2019*b*), 'Should we target jobs at distressed places, and if so, how?'.

Bartik, T. J. and Sotherland, N. (2019), 'Local job multipliers in the United States: Variation with local characteristics and with high-tech shocks'.

Bayer, P., McMillan, R., Murphy, A. and Timmins, C. (2016), 'A dynamic model of demand for houses and neighborhoods', *Econometrica* **84**(3), 893–942.

Bennett, D. L. (2019), 'Infrastructure investments and entrepreneurial dynamism in the U.S.', *Journal of Business Venturing* **34**(5), 105907.

Blanchard, O. and Katz, L. F. (1992), 'Regional evolutions', *Brookings Papers on Economic Activity* **1992**(1), 1–75.

Bloom, N., Jones, C. I., Van Reenen, J. and Webb, M. (2017), Are ideas getting harder to find?, Technical report, National Bureau of Economic Research.

Boehmke, F. J., Brockway, M., Desmarais, B. A., Harden, J. J., LaCombe, S., Linder, F. and Wallach, H. (2020), 'Spid: A new database for inferring public policy innovativeness and diffusion networks', *Policy Studies Journal* **48**(2), 517–545.

Borusyak, K., Hull, P. and Jaravel, X. (2018), Quasi-experimental shift-share research designs, Technical report, National Bureau of Economic Research.

Bremmer, D. S. and Kesselring, R. G. (1993), 'The determinants of regional manufacturing investment: a simultaneous equations approach', *Journal of Regional Analysis and Policy* **23**(1100-2016-89700), 3–27.

Broxterman, D. A. and Larson, W. D. (2020), 'An empirical examination of shift-share instruments', *Journal of Regional Science* .

Canay, I. A. (2011), 'A simple approach to quantile regression for panel data', *The Econometrics Journal* **14**(3), 368–386.

Cao, D., Hyatt, H., Mukoyama, T. and Sager, E. (2020), Firm growth through new establishments, Technical report, Georgetown University, Department of Economics.

Capello, R. and Nijkamp, P. (2010), *Handbook of regional growth and development theories*, Edward Elgar Publishing.

Chen, L. and Huo, Y. (2019), 'A simple estimator for quantile panel data models using smoothed quantile regressions', *The Econometrics Journal* .

Chernozhukov, V. and Hansen, C. (2008), 'Instrumental variable quantile regression: A robust inference approach', *Journal of Econometrics* **142**(1), 379–398.

Cohen, L., Coval, J. and Malloy, C. (2011), 'Do powerful politicians cause corporate downsizing?', *Journal of Political Economy* **119**(6), 1015–1060.

Cohen, W. M. and Levinthal, D. A. (1990), 'Absorptive capacity: A new perspective on learning and innovation', *Administrative science quarterly* pp. 128–152.

Colella, F., Lalive, R., Sakalli, S. O. and Thoenig, M. (2019), 'Inference with arbitrary clustering, IZA Discussion Paper n. 12584'.

Crouzet, N. and Eberly, J. (2018), Intangibles, investment, and efficiency, *in* 'AEA Papers and Proceedings', Vol. 108, pp. 426–31.

Dalenberg, D. R. and Partridge, M. D. (1997), 'Public infrastructure and wages: Public capital's role as a productive input and household amenity', *Land Economics* pp. 268–284.

Davis, S. J. and Haltiwanger, J. (2014), Labor market fluidity and economic performance, Technical report, National Bureau of Economic Research.

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

Decker, R. A., Haltiwanger, J., Jarmin, R. S. and Miranda, J. (2016*a*), 'Declining business dynamism: What we know and the way forward', *American Economic Review* **106**(5), 203–07.

Decker, R. A., Haltiwanger, J., Jarmin, R. S. and Miranda, J. (2016*b*), 'Where has all the skewness gone? The decline in high-growth (young) firms in the U.S.', *European Economic Review* **86**, 4–23.

Decker, R., Haltiwanger, J., Jarmin, R. and Miranda, J. (2014), 'The role of entrepreneurship in US job creation and economic dynamism', *Journal of Economic Perspectives* **28**(3), 3–24.

Dorn, D., Hanson, G. H., Pisano, G., Shu, P. et al. (2020), 'Foreign competition and domestic innovation: Evidence from us patents', *American Economic Review: Insights* **2**(3), 357–74.

Dupor, B. (2017), 'So, why didn't the 2009 recovery act improve the nation's highways and bridges?'.

Eckert, F., Fort, T. C., Schott, P. K. and Yang, N. J. (2020), Imputing Missing Values in the US Census Bureau's County Business Patterns, Technical report, National Bureau of Economic Research.

Elhorst, J. P. (2021), 'The dynamic general nesting spatial econometric model for spatial panels with common factors: Further raising the bar', *Review of Regional Research* pp. 1–19.

Feldman, M. P. and Florida, R. (1994), 'The geographic sources of innovation: technological infrastructure and product innovation in the United States', *Annals of the association of American Geographers* **84**(2), 210–229.

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

Galvao, A. F., Montes-Rojas, G., Sosa-Escudero, W. and Wang, L. (2013), 'Tests for skewness and kurtosis in the one-way error component model', *Journal of Multivariate Analysis* **122**, 35–52.

Garin, A. (2019), 'Putting america to work, where? evidence on the effectiveness of infrastructure construction as a locally targeted employment policy', *Journal of Urban Economics* **111**, 108–131.

Gibbons, S. and Overman, H. G. (2012), 'Mostly pointless spatial econometrics?', *Journal of regional Science* **52**(2), 172–191.

Goldsmith-Pinkham, P., Sorkin, I. and Swift, H. (2018), Bartik instruments: What, when, why, and how, Technical report, National Bureau of Economic Research.

Goolsbee, A. (1998), 'Does government R&D policy mainly benefit scientists and engineers', *The American Economic Review* **88**(2), 298–302.

Gordon, R. J. (2012), Is US economic growth over? Faltering innovation confronts the six headwinds, Technical report, National Bureau of Economic Research.

Greenaway-McGrevy, R., Han, C. and Sul, D. (2012), 'Estimating the number of common factors in serially dependent approximate factor models', *Economics Letters* **116**(3), 531–534.

Greenaway-McGrevy, R. and Hood, K. (2019), 'Aggregate effects and measuring regional dynamics', *Papers in Regional Science* **98**(5), 1955–1991.

Gunnella, V., Mastromarco, C., Serlenga, L. and Shin, Y. (2015), 'The euro effects on intra-eu trade flows and balance: evidence from the cross sectionally dependent panel gravity models'.

Gyourko, J. and Tracy, J. (1989), 'The importance of local fiscal conditions in analyzing local labor markets', *Journal of Political economy* **97**(5), 1208–1231.

Hall, R. E. (2018), New evidence on the markup of prices over marginal costs and the role of mega-firms in the us economy, Technical report, National Bureau of Economic Research.

Halleck Vega, S. and Elhorst, J. P. (2015), 'The SLX model', *Journal of Regional Science* **55**(3), 339–363.

Haltiwanger, J., Hathaway, I. and Miranda, J. (2014), 'Declining business dynamism in the U.S. high-technology sector', *Available at SSRN 2397310* .

Haltiwanger, J., Jarmin, R. S., Kulick, R. B. and Miranda, J. (2016), 'High growth young firms: Contribution to job, output and productivity growth', *US Census Bureau Center for Economic Studies Paper No. CES-WP-16-49* .

Harding, M. and Lamarche, C. (2009), 'A quantile regression approach for estimating panel data models using instrumental variables', *Economics Letters* **104**(3), 133–135.

Hooper, E., Peters, S. and Pintus, P. (2018), 'The causal effect of infrastructure investments on income inequality: Evidence from US states'.

Imbens, G. W. and Kolesar, M. (2016), 'Robust standard errors in small samples: Some practical advice', *Review of Economics and Statistics* **98**(4), 701–712.

Jacobs, J. (1969), *The economy of cities*, Vintage.

Jaffe, A. B. (1989), 'Real effects of academic research', *The American economic review* pp. 957–970.

Juodis, A. and Reese, S. (2021), 'The incidental parameters problem in testing for remaining cross-section correlation', *Journal of Business & Economic Statistics* pp. 1–13.

Karabarbounis, L. and Neiman, B. (2014), 'The global decline of the labor share', *The Quarterly journal of economics* **129**(1), 61–103.

Kehrig, M. and Vincent, N. (2018), The micro-level anatomy of the labor share decline, Technical report, National Bureau of Economic Research.

Kemeny, T. and Osman, T. (2018), 'The wider impacts of high-technology employment: Evidence from US cities', *Research Policy* **47**(9), 1729–1740.

Khandker, S. R. and Koolwal, G. B. (2010), 'How infrastructure and financial institutions affect rural income and poverty: evidence from bangladesh', *The journal of development studies* **46**(6), 1109–1137.

Khandker, S. R. and Koolwal, G. B. (2011), *Estimating the long-term impacts of rural roads: a dynamic panel approach*, The World Bank.

Kleinknecht, A. (2020), 'The (negative) impact of supply-side labour market reforms on productivity: an overview of the evidence', *Cambridge Journal of Economics* .

Koenker, R. (2004), 'Quantile regression for longitudinal data', *Journal of Multivariate Analysis* **91**(1), 74–89.

Leduc, S. and Wilson, D. (2013), 'Roads to prosperity or bridges to nowhere? theory and evidence on the impact of public infrastructure investment', *NBER Macroeconomics Annual* **27**(1), 89–142.

LeSage, J. and Pace, R. K. (2009), *Introduction to spatial econometrics*, Chapman and Hall/CRC.

Ljungqvist, A. and Smolyansky, M. (2014), To cut or not to cut? on the impact of corporate taxes on employment and income, Technical report, National Bureau of Economic Research.

Miller, J. B. (2013), *Principles of public and private infrastructure delivery*, Vol. 101, Springer Science & Business Media.

Moran, S., Kierzek, R. and Turner, D. H. (1993), 'Binding of guanosine and 3'splice site analogs to a group i ribozyme: Interactions with functional groups of guanosine and with additional nucleotides', *Biochemistry* **32**(19), 5247–5256.

Moretti, E. (2010), 'Local multipliers', *American Economic Review* **100**(2), 373–77.

Moretti, E. (2013), 'Real wage inequality', *American Economic Journal: Applied Economics* **5**(1), 65–103.

Moretti, E. and Thulin, P. (2013), 'Local multipliers and human capital in the United States and Sweden', *Industrial and Corporate Change* **22**(1), 339–362.

Moretti, E. and Wilson, D. J. (2014), 'State incentives for innovation, star scientists and jobs: Evidence from biotech', *Journal of Urban Economics* **79**, 20–38.

Moretti, E. and Wilson, D. J. (2017), 'The effect of state taxes on the geographical location of top earners: Evidence from star scientists', *American Economic Review* **107**(7), 1858–1903.

Pereira, A. M. and Andraz, J. M. (2003), 'On the impact of public investment on the performance of US industries', *Public Finance Review* **31**(1), 66–90.

Pereira, A. M. and Andraz, J. M. (2004), 'Public highway spending and state spillovers in the USA', *Applied Economics Letters* **11**(12), 785–788.

Pereira, A. M. and Andraz, J. M. (2012*a*), 'On the effects of highway investment on the regional concentration of economic activity in the USA', *Portuguese Economic Journal* **11**(3), 165–170.

Pereira, A. M. and Andraz, J. M. (2012*b*), 'On the regional incidence of highway investments in the USA', *The Annals of regional science* **48**(3), 819–838.

Pesaran, M. H. (2006), 'Estimation and inference in large heterogeneous panels with a multifactor error structure', *Econometrica* **74**(4), 967–1012.

Petralia, S., Balland, P.-A. and Rigby, D. L. (2016), 'Unveiling the geography of historical patents in the united states from 1836 to 1975', *Scientific data* **3**(1), 1–14.

Pierson, K., Hand, M. L. and Thompson, F. (2015), 'The government finance database: A common resource for quantitative research in public financial analysis', *PloS one* **10**(6).

Powell, D. (2016), 'Quantile regression with nonadditive fixed effects', *Quantile Treatment Effects* pp. 1–28.

Ramey, V. A. (2020), The macroeconomic consequences of infrastructure investment, Technical report, National Bureau of Economic Research.

Rauch, J. E. (1993), 'Does history matter only when it matters little? the case of city-industry location', *The quarterly journal of economics* **108**(3), 843–867.

Roback, J. (1982), 'Wages, rents, and the quality of life', *Journal of political Economy* **90**(6), 1257–1278.

Santoleri, P. et al. (2019), 'Innovation and job creation in (high-growth) new firms', *Industrial and Corporate Change* .

Schumpeter, J. (1934), 'The theory of economic development Harvard University Press', *Cambridge, MA* .

Schumpeter, J. A. (1942), 'Capitalism, socialism and democracy', *New York* .

Serrato, J. C. S. and Zidar, O. (2018), 'The structure of state corporate taxation and its impact on state tax revenues and economic activity', *Journal of Public Economics* **167**, 158–176.

Stansel, D. and McMahon, F. (2018), *Economic Freedom of North America 2013*, Fraser Institute Vancouver.

Stewart, C. and Nelson, G. (2005), 'Garrison Nelson, Committees in the US Congress, 1947-1992, data'.

Stewart, C. and Woon, J. (2017), 'Congressional committee assignments, 103rd to 114th congresses, 1993–2017', *House and Senate* .

Trogen, P. (1999), 'Which economic development policies work: determinants of state per capita income', *International Journal of Economic Development* **1**(3), 256–279.

Tuszynski, M. P. and Stansel, D. (2018), 'Targeted state economic development incentives and entrepreneurship', *Journal of Entrepreneurship and Public Policy* .

Vaghul, K. and Zipperer, B. (2016), 'Historical state and sub-state minimum wage data', *Washington Center for Equitable Growth* .

Walsh, C. (2019), 'Firm creation and local growth', *Available at SSRN 3496782* .

Williamson, O. E. (1998), 'The institutions of governance', *The American Economic Review* **88**(2), 75–79.

Wilson, D. J. (2009), 'Beggar thy neighbor? The in-state, out-of-state, and aggregate effects of R&D tax credits', *The Review of Economics and Statistics* **91**(2), 431–436.

Zidar, O. (2019), 'Tax cuts for whom? heterogeneous effects of income tax changes on growth and employment', *Journal of Political Economy* **127**(3), 1437–1472.

Figure 1: A simple model of regional employment with incentives

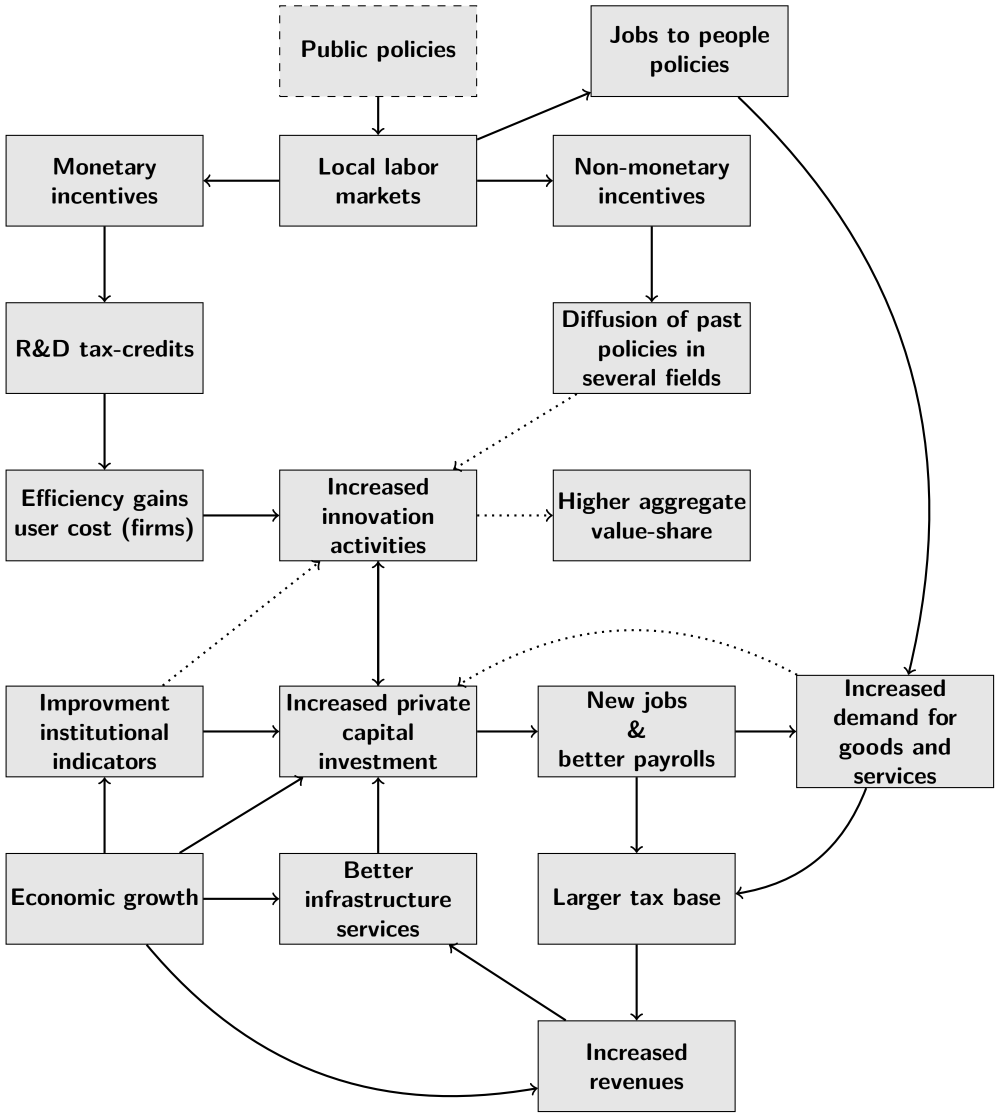
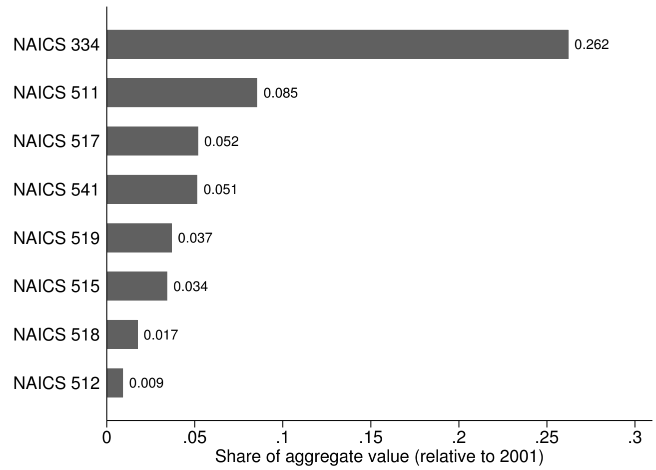


Figure 2: Average aggregate value share within high-tech industries

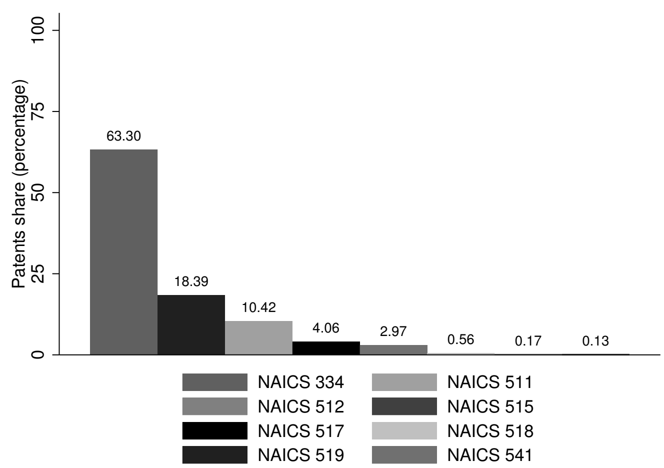
U.S public companies by subsector: 1990-2015



Note: This figure presents the aggregate value share (AV) for U.S public firms in 2001. In doing so, we follow Crouzet and Eberly (2018), Hall (2018) methodology. High-tech industries (software and telecommunications) are collapsed at three-digit North American Industry Classification Structure (NAICS) and KLEMS level: 334, 511, 512, 515, 517, 518, 519 and 541 respectively. Source: author own calculations.

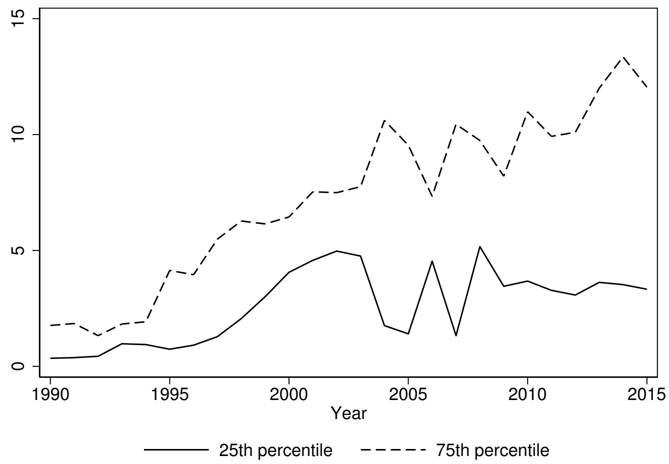
Figure 3: Share of patents U.S high-tech industries

U.S public companies by subsector: 1990-2015



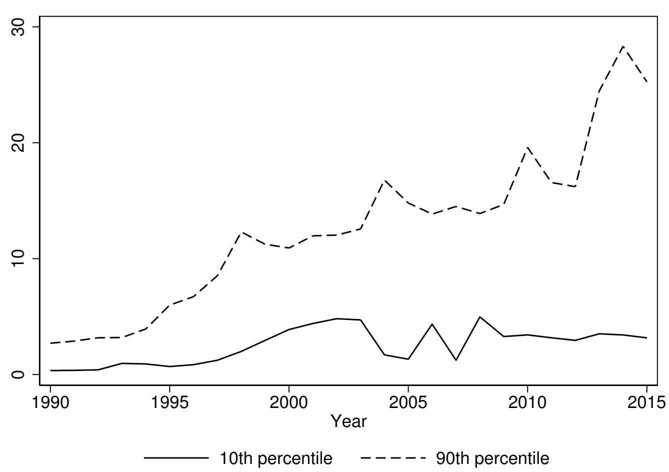
Note: This figure presents the share of patents within high-tech sectors. Patents data comes from patents view, a crosswalk of U.S public firms (COMPUSTAT) Dorn et al. (2020) and Arora et al. (2017). High-tech industries (software and telecommunications) are collapsed at three-digit North American Industry Classification Structure (NAICS) and KLEMS level: 334, 511, 512, 515, 517, 518, 519 and 541 respectively. Source: author own calculations.

Figure 4: Distribution of innovation outcomes by quartiles



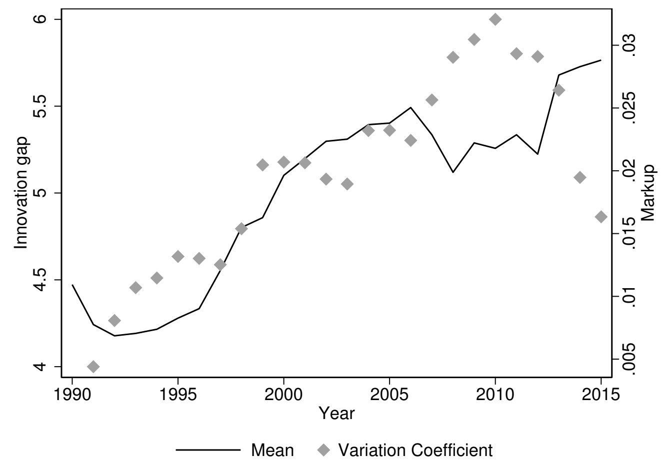
Note: Patents data comes from patents view, a crosswalk of U.S public firms (COMPUSTAT) Dorn et al. (2020) and Arora et al. (2017). The quartiles of per capita patents are constructed as the weighted sum of aggregate value share (AV) of U.S public firms in 2001 (Crouzet and Eberly, 2018, Hall, 2018). High-tech industries (software and telecommunications) are collapsed at three-digit North American Industry Classification Structure (NAICS) and KLEMS level: 334, 511, 512, 515, 517, 518, 519 and 541 respectively. For a clear visual inspection, patents are rescaled per 100,000 population and transformed into logarithms. The analysis excludes States with missing data (West Virginia, Wyoming), the District of Columbia (DC) and States not part of the contiguous United States (Alaska and Hawaii). Source: author own calculations.

Figure 5: Distribution of innovation outcomes: top and bottom percentiles



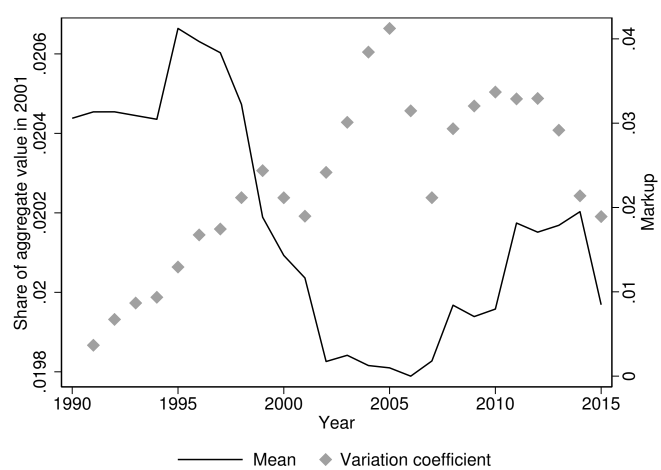
Note: Patents data comes from patents view, a crosswalk of U.S public firms (COMPUSTAT) Dorn et al. (2020) and Arora et al. (2017). The percentiles of per capita patents are constructed as the weighted sum of aggregate value share (AV) of U.S public firms in 2001 (Crouzet and Eberly, 2018, Hall, 2018). High-tech industries (software and telecommunications) are collapsed at three-digit North American Industry Classification Structure (NAICS) and KLEMS level: 334, 511, 512, 515, 517, 518, 519 and 541 respectively. For a clear visual inspection, patents are rescaled per 100,000 population and transformed into logarithms. The analysis excludes States with missing data (West Virginia, Wyoming), the District of Columbia (DC) and States not part of the contiguous United States (Alaska and Hawaii). Source: author own calculations.

Figure 6: Innovation gap and market concentration: 1990-2015



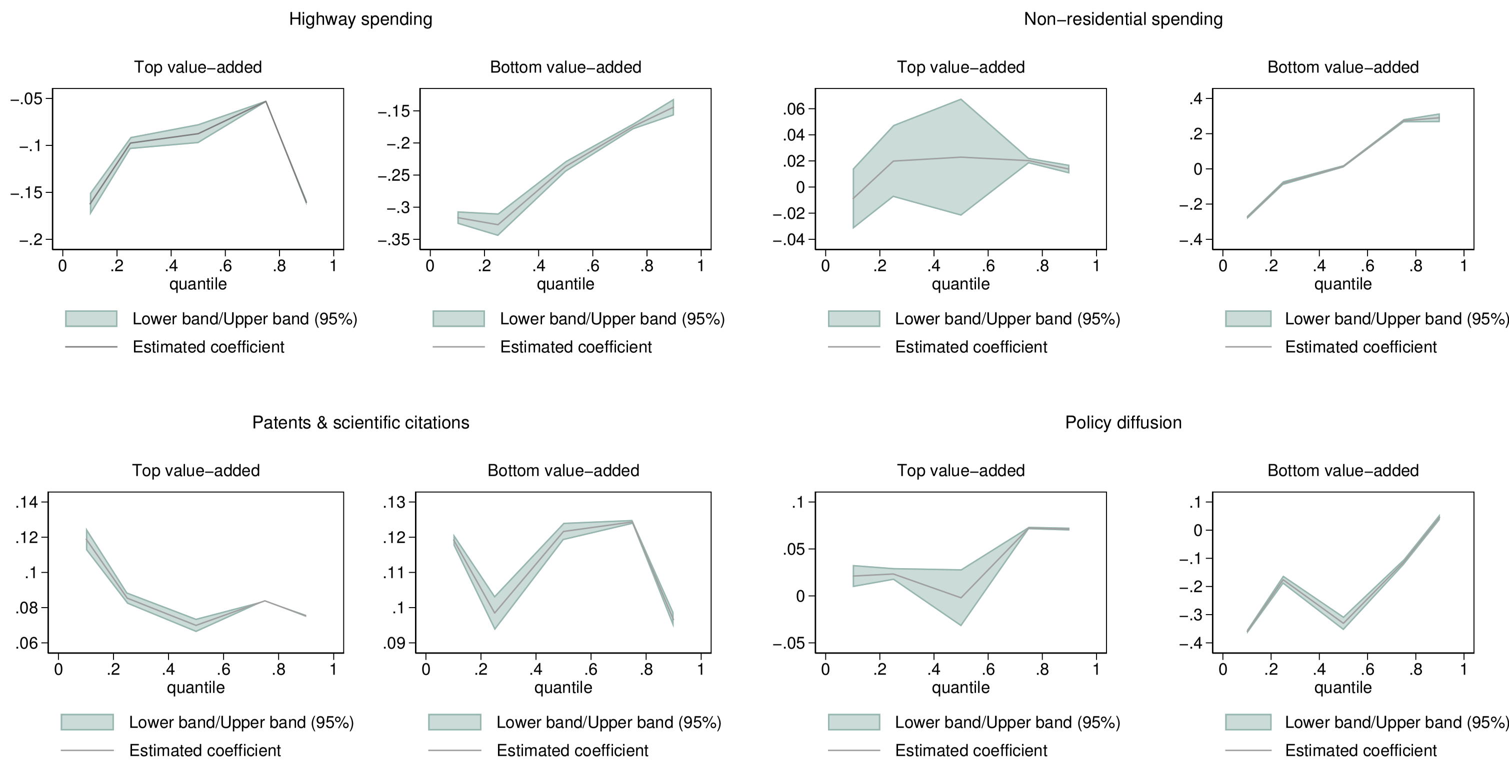
Note: Innovation gap (line) is the difference between the top and bottom percentiles of patents and scientific citations (i.e., 90^{th} and 10^{th}) whereas the markup (dots) of public U.S firms (COMPUSTAT) was calculated following Crouzet and Eberly (2018) and Hall (2018) methodology. High-tech industries (software and telecommunications) are collapsed at three-digit North American Industry Classification Structure (NAICS) and KLEMS level: 334, 511, 512, 515, 517, 518, 519 and 541 respectively. The analysis excludes States with missing data (West Virginia, Wyoming), the District of Columbia (DC) as well as States not part of the contiguous United States (Alaska and Hawaii). To avoid business cycles effects, rolling windows are included. Source: author own calculations.

Figure 7: Share of aggregate value and market concentration: 1990-2015



Note: Aggregate value share in 2001 (line) and markup (dots) of public U.S firms (COMPUSTAT) were calculated following Crouzet and Eberly (2018) and Hall (2018) methodology. High-tech industries (software and telecommunications) are collapsed at three-digit North American Industry Classification Structure (NAICS) and KLEMS level: 334, 511, 512, 515, 517, 518, 519 and 541 respectively. To avoid business cycles effects, rolling windows are included. The analysis excludes States with missing data (West Virginia, Wyoming), the District of Columbia (DC) as well as States not part of the contiguous United States (Alaska and Hawaii). Source: author own calculations.

Figure 8: Wages, infrastructure and innovation outcomes: Quantile panel regression with nonadditive fixed effects



Note: The quantile panel model with nonadditive fixed effects is estimated including the GDP growth rate (excluding both measures of infrastructure) and dummies by U.S Census division and decades. Robust standard errors in brackets are computed by bootstrap method with 800 replications using the Metropolis-within-Gibbs sampler. *** p¡0.01, ** p¡0.05, * p¡0.1. Source: author own calculations.

Table 1: Summary Statistics

Variable	Mean	SD	Min	Max
Highway spending	269.100	128.100	65.690	1,081
Private non-residential spending	88.450	47.100	11.850	505.600
(log of) Employment (top value-added)	11.690	1.100	9.259	14.320
(log of) Employment (bottom value-added)	9.491	1.070	7.016	12.640
(log of) Real wage (top value-added)	10.720	0.941	8.195	13.080
(log of) Real wage (bottom value-added)	10.380	0.890	8.125	12.810
(log of) Real GDP per capita	10.740	0.191	10.250	11.390
(growth rate) RGDP per capita (net infrastructure)	1.215	3.454	-16.440	14.570
Population density (100 s persons per sqmi)	1.901	2.555	0.048	12.180
Unemployment rate	5.582	1.888	2.300	13.700
High-school graduation rate	0.627	0.043	0.499	0.748
College graduation rate	0.178	0.042	0.079	0.306
Patents per capita (all high-tech)	6.908	17.190	0.000	173.800
Patents per capita (10^{th} percentile)	2.843	14.290	0.000	173.800
Patents per capita (25^{th} percentile)	2.918	14.280	0.000	173.800
Patents per capita (75^{th} percentile)	8.001	23.030	0.000	276.100
Patents per capita (90^{th} percentile)	14.010	31.610	0.000	333.100
Share of aggregate value (all high-tech)	0.020	0.002	0.012	0.029
Average markup (all high-tech)	1.828	0.210	0.862	2.005
Policy diffusion score	0.047	0.022	0.000	0.175
R&D user cost	1.144	0.022	1.105	1.202
Government spending score	6.684	1.535	2.781	9.701
Tax distortion score	5.660	0.890	2.560	7.760
Labor market friction score	5.485	1.074	1.771	8.670
Economic freedom score	5.944	0.928	2.967	8.046
Observations		1104		

Notes:

a. Real capital outlays of highway spending includes: maintenance, operation, purchases of equipment, toll highways, bridges, tunnels, ferries, street lighting, snow and ice removal. The real measure was computed using the State price index provided by the Bureau of Economic Analysis (BEA) (table 3.9.4) and normalized by state population to control for business cycle effects. Real private non-residential spending was obtained from U.S Census and deflated by the BLS consumer price index CPI-U (2012=100) and normalized by state population to control for business cycle effects.

b. Job-creation rates (private employment) are calculated as the difference of the ratio in annual employment levels in period t with respect to t-1 times 100 and nominal wage bills (i.e. annual payrolls) were deflated using the BLS consumer price index CPI-U (2012=100) and rescaled into logarithms. Source: Quarterly Census of Employment & Wages (QCEW).

c. Population density is built dividing the land area (Sq. mi.) of the U.S Census by state population while annual unemployment rate data comes from the BLS civilian civilian noninstitutional population.

d. Education attainment (per-capita) data comes from Frank (2009), unemployment rates are taken from the Bureau Labor Statistics.

e. We concentrate on the most innovative high-tech industries (relative to aggregate value share) of patents and scientific citations of all software and telecommunications industries). Thus, our top 4 group, based on the three-digit North American Industry Classification Structure (NAICS) includes the 334, 511, 517 and 541 while the bottom 4 group encompass 512, 515, 518 and 519 sectors.

f. Patents data comes from Arora et al. (2017), patents view and a crosswalk of U.S public firms (COMPUSTAT) Dorn et al. (2020). The percentiles of per capita patents are constructed as the weighted sum of aggregate value share (AV) of U.S public firms in 2001. In lieu with empirical literature, patents are rescaled per 100,000 population. Average markups and aggregate 2001 value of public U.S firms were calculated following Crouzet and Eberly (2018) and Hall (2018) methodology.

g. The policy diffusion score includes 204 policies for the following sectors: education, energy, environment, domestic commerce, housing, labor and macroeconomics (Boehmke et al., 2020). To avoid business cycles rolling windows are included.

h. R&D user cost was constructed following Wilson (2009). For the expenditure share (s) we use COMPUSTAT data (instead of IRS) which is on average 0.1243 between the years 1965 and 2015 rather than 0.5. Particularly, we concentrate on the upper bound (90th percentile) which is about half of the IRS income data (i.e. 0.2588).

i. The District of Columbia (DC) as well as Alaska and Hawaii (not part of the contiguous United States) are excluded from the analysis. Source: author own calculations.

Table 2: Correlations: Infrastructure, employment, innovation and wages

	Highway spending		Private non-residential spending		Innovation	
	Panel A					
Census Division	Top value-added	Bottom value-added	Top value-added	Bottom value-added	Top value-added	Bottom value-added
New England	0.061	−0.101	0.024	0.185*	0.199**	0.020
Middle Atlantic	−0.156	−0.111	0.177	0.064	0.111	−0.130
East North	−0.170*	−0.122	0.117	−0.113	0.067	0.124
West North	0.069	−0.122	0.137	−0.133	0.146*	−0.118
South Atlantic	−0.058	−0.256**	0.205**	0.169**	0.026	0.015
East South	−0.118	−0.126	0.123	0.012	−0.138	0.068
West South	−0.175	−0.233**	0.057	−0.059	−0.101	−0.224**
Mountain	−0.094	−0.172**	0.199**	0.105	0.034	−0.085
Pacific	0.100	0.052	0.328**	0.051	0.141	0.445***
All Divisions	−0.029	−0.146***	0.154***	0.076**	0.063*	0.003
	Panel B					
Census Division	Top value-added	Bottom value-added	Top value-added	Bottom value-added	Top value-added	Bottom value-added
New England	−0.386***	−0.444***	−0.481***	−0.523***	0.003	0.057
Middle Atlantic	0.369**	0.218*	−0.203*	−0.127	0.372**	0.307**
East North	0.129	0.148	−0.156	−0.270**	0.354***	0.456***
West North	0.715***	0.653***	0.273**	0.215**	0.388***	0.338***
South Atlantic	0.348***	0.346***	−0.451***	−0.456***	−0.111	−0.086
East South	0.060	−0.050	−0.306**	−0.248**	0.539***	0.487***
West South	0.275**	0.306**	−0.205**	−0.182	−0.166	−0.181
Mountain	0.645***	0.638***	−0.253**	−0.219**	−0.030	−0.040
Pacific	0.649***	0.540***	0.349**	0.222*	0.407**	0.404**
All Divisions	0.395***	0.382***	−0.020	−0.037	0.085**	0.090**
	Panel C					
Census Division	P25th	P75th	P25th	P75th	P25th	P75th
New England	−0.094	0.023	−0.449***	−0.142	0.699***	0.164*
Middle Atlantic	0.053	−0.337**	0.084	0.483***	0.335***	0.182
East North	0.211**	0.529***	0.229**	0.224**	0.478***	0.033
West North	0.563***	0.090	0.239**	0.052	0.654***	0.306**
South Atlantic	0.364***	0.428***	−0.304***	−0.338***	0.790***	0.846***
East South	0.776***	0.290**	−0.261**	−0.223*	0.283**	0.283**
West South	0.278**	0.030	−0.613***	−0.239*	0.954***	0.306**
Mountain	0.321***	0.234**	−0.153**	−0.116	0.609***	0.481***
Pacific	0.592***	0.565***	0.380**	0.138	0.935***	0.751***
All Divisions	0.355***	0.250***	−0.010	−0.004	0.756***	0.605***
	Panel D					
Census Division	P10th	P90th	P10th	P90th	P10th	P90th
New England	0.107	0.223	−0.288**	0.330***	0.289**	−0.738***
Middle Atlantic	−0.183	−0.458***	0.197	0.432***	0.008	0.063
East North	−0.086	0.315**	−0.205**	−0.085	0.045	−0.352***
West North	0.264**	−0.629***	0.205**	−0.235**	0.249**	−0.449***
South Atlantic	0.074	0.145*	−0.337**	−0.331**	0.158**	0.286
East South	0.531***	0.142	−0.121	−0.050	0.198	0.107
West South	0.278**	−0.283**	−0.241*	0.569***	0.491***	−0.934***
Mountain	0.195**	−0.088	−0.160*	−0.002	0.503***	−0.059
Pacific	0.350**	−0.115	−0.109	−0.381**	0.474***	−0.160**
All Divisions	0.197***	−0.163***	−0.046	0.038	0.333**	−0.237**

Notes:

a. Panels A and B show the relation among infrastructure, employment, innovation outcomes and wage bills for high-tech industries with different aggregate value share. Panels C and D consider only the relationship between infrastructure expenditures, high-tech wages and conditional percentiles of innovation outcomes. For statistical convenience, all variables are rescaled into logarithms.

b. Real capital outlays of highway spending includes: maintenance, operation, purchases of equipment, toll highways, bridges, tunnels, ferries, street lighting, snow and ice removal. The real measure was computed using the State price index provided by the Bureau of Economic Analysis (BEA) (table 3.9.4), rescaled by state population to control for business cycle effects. Real private non-residential spending was obtained from U.S Census and deflated by the BLS consumer price index CPI-U (2012=100) rescaled by state population to control for business cycle effects.

c. Job-creation rates (private employment) were calculated as the difference of the ratio in employment levels in period t with respect to t-1 times 100. Nominal wage bills (i.e. annual payrolls) were deflated by the BLS consumer price index CPI-U (2012=100) rescaled by state population. Source: Quarterly Census of Employment & Wages (QCEW).

d. We concentrate on the most innovative high-tech industries (relative to aggregate value share) of patents and scientific citations of all software and telecommunications industries). Thus, our top 4 group, based on the three-digit North American Industry Classification Structure (NAICS) includes the 334, 511, 517 and 541 while the bottom 4 group encompass 512, 515, 518 and 519 sectors.

e. For innovation outcomes (third column) we consider two measures. Panels A and B regard as non-monetary incentives the diffusion of public policies (score) in the fields of: macroeconomics, labor, education, environment, energy, housing and domestic commerce (Boehmke et al., 2020). Panels C and D use the 25th, 75th, 10th and 90th percentiles of per capita (normalized per 100,000 population) patents from U.S public firms (COMPUSTAT) weighted by their respective value share in 2001 (Crouzet and Eberly (2018), Hall, 2018).

f. States with missing data (West Virginia, Wyoming), the District of Columbia (DC) and States not part of the contiguous United States (Alaska and Hawaii) are excluded. ***, ** and * indicate significance at the 1%, 5% and 10% levels. Source: author own calculations.

Table 3: Infrastructure, employment and incentives: GMM Estimates

dv: job creation rate	All			Top value-added			Bottom value-added		
	(I)	(II)	(III)	(I)	(II)	(III)	(I)	(II)	(III)
(log of) Real highway spending	−0.241	−0.270	−1.019**	−0.194	−0.213	−1.001**	−2.285**	−2.279**	−2.025**
	[0.341]	[0.355]	[0.460]	[0.351]	[0.364]	[0.476]	[0.663]	[0.675]	[0.878]
(log of) Real non-residential spending	0.740**	0.724**	0.939**	0.888**	0.881**	1.064**	0.545	0.416	0.762
	[0.275]	[0.276]	[0.395]	[0.278]	[0.280]	[0.421]	[0.588]	[0.585]	[0.734]
(log of) Policy diffusion score	0.799**	0.809**	0.774**	0.893**	0.896**	0.777**	0.366	0.442	0.692
	[0.296]	[0.296]	[0.321]	[0.305]	[0.304]	[0.330]	[0.493]	[0.494]	[0.651]
R&D incentive (0-3 years)			36.013**			27.736*			61.375**
			[15.878]			[16.791]			[21.741]
R&D user cost (0-3 years)			−6.346			−7.155			−15.769**
			[4.674]			[4.918]			[7.575]
High-tech incentive × R&D user cost			−10.570**			−8.166*			−17.889**
			[4.647]			[4.912]			[6.334]
GDP growth rate (net infrastructure)	0.175***	0.176***	0.118***	0.175***	0.175***	0.117**	0.152**	0.160**	0.089
	[0.032]	[0.032]	[0.033]	[0.033]	[0.033]	[0.035]	[0.074]	[0.074]	[0.066]
Hansen J test (p-value)	(0.482)	(0.512)	(0.275)	(0.649)	(0.661)	(0.374)	(0.815)	(0.830)	(0.496)
Underidentification test (p-value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Exogeneity test (p-value)	(0.837)	(0.835)	(0.710)	(0.891)	(0.890)	(0.829)	(0.581)	(0.594)	(0.335)
AR Weak IV test (p-value)	(0.556)	(0.577)	(0.107)	(0.732)	(0.738)	(0.167)	(0.061)	(0.070)	(0.173)
CSD_W	0.256	2.976	−2.803	0.755	−1.333	2.243	0.344	2.877	1.504
R-squared (uncentered)	0.218	0.219	0.435	0.216	0.216	0.420	0.067	0.069	0.249
Number of states	48	48	48	48	48	48	48	48	46
Number of observations	1007	1007	1007	992	992	992	893	893	867
Geography controls	N	Y	Y	N	Y	Y	N	Y	Y
Socioeconomic controls	N	N	Y	N	N	Y	N	N	Y
Institutional controls	N	N	Y	N	N	Y	N	N	Y
Supply controls	N	N	Y	N	N	Y	N	N	Y

Notes:

a. All models include Census divisions fixed effects. Additionally, since annual public expenditures vary in sample size, I include a dummy break to account for time fixed effects without compromising the number of covariates. Indeed, its inclusion (not reported) was highly significant in almost all specifications, meaning that this effect cannot be disregarded. Geography controls include average temperatures in winter (January) and summer (July) while socioeconomic ones regard the high school and college graduation rates from total state population. Last but not least, we account for the institutional quality by employing all the disaggregated scores (i.e., subcategories) of the EFNA named: Government spending, Taxation and Labor Market Freedom. Finally, we control for supply shocks with a dummy variable equal to one when the economy had downturns periods (e.g., 1990-91, 2001 and 2008-2010) and zero in contrary case.

b. Real capital outlays of highway spending includes: maintenance, operation, purchases of equipment, toll highways, bridges, tunnels, ferries, street lighting, snow and ice removal. The real measure was computed using the State price index provided by the Bureau of Economic Analysis (BEA) (table 3.9.4) and rescaled into logarithms. Real private non-residential spending comes from U.S Census and deflated by the BLS consumer price index CPI-U (2012=100) and rescaled by state population to control for business cycle effects. Data starts from 1993 onwards.

c. Job-creation rates (private employment) were calculated as the difference of the ratio in employment levels in period t with respect to t-1 times 100. Source: Quarterly Census of Employment & Wages (QCEW).

d. We concentrate on the most innovative high-tech industries (relative to aggregate value share) of patents and scientific citations of all software and telecommunications industries). Thus, our top 4 group, based on the three-digit North American Industry Classification Structure (NAICS) includes the 334, 511, 517 and 541 while the bottom 4 group encompass 512, 515, 518 and 519 sectors.

e. For the high-tech incentive dummy we consider those states in which R&D tax-credits were positive. Data comes from the Panel Database of Incentives and Taxes (PDIT) Bartik (2017). While R&D user cost was constructed following Wilson (2009). For the expenditure share (s) we use COMPUSTAT data (instead of IRS) which is on average 0.1243 between the years 1965 and 2015 rather than 0.5. Particularly, we concentrate on the upper bound (90^{th} percentile) which is about half of the IRS income data (i.e., 0.2588).

f. The policy diffusion score includes 204 policies for the following sectors: education, energy, environment, domestic commerce, housing, labor and macroeconomics (Boehmke et al., 2020).

g. The total number of draws of the weighted CSD statistic were set to 10 to avoid distortions from lower order terms. In this way, the statistic can be constructed accounting for both serial correlation and CSD (Juodis and Reese, 2021). Autocorrelated standard errors (HAC) with the Newey-West bandwidth adjustment in brackets small sample adjustment. The number of optimal lags for the bandwidth is set to 3.

h. The District of Columbia (DC) as well as Alaska and Hawaii (not part of the contiguous United States) were excluded from the analysis. Also we account for outliers (Delaware, Maine). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: author own calculations.

Table 4: Infrastructure, employment, incentives and innovation: GMM Estimates

dv: job creation rate	All			Top value-added			Bottom value-added		
	(I)	(II)	(III)	(I)	(II)	(III)	(I)	(II)	(III)
(log of) Real highway spending	−0.780	−0.792	−1.490**	−0.644	−0.647	−1.399**	−2.719**	−2.748**	−2.846**
	[0.497]	[0.497]	[0.541]	[0.513]	[0.514]	[0.556]	[0.898]	[0.895]	[0.998]
(log of) Real non-residential spending	1.037**	0.982**	1.220**	1.138**	1.098**	1.338**	0.316	0.216	0.847
	[0.339]	[0.341]	[0.477]	[0.367]	[0.370]	[0.513]	[0.655]	[0.655]	[0.834]
(log of) Patents (10^{th} percentile)	0.044	0.036	0.113	0.004	−0.000	0.095	0.127	0.109	0.177
	[0.114]	[0.114]	[0.096]	[0.113]	[0.112]	[0.097]	[0.209]	[0.211]	[0.191]
(log of) Patents (25^{th} percentile)	−0.540*	−0.522*	−0.463	−0.385	−0.376	−0.422	−0.836*	−0.812	−0.651
	[0.312]	[0.311]	[0.315]	[0.327]	[0.326]	[0.320]	[0.506]	[0.509]	[0.541]
(log of) Patents (75^{th} percentile)	0.490*	0.483*	0.518**	0.372	0.366	0.419*	0.800*	0.782*	1.191**
	[0.256]	[0.259]	[0.245]	[0.247]	[0.252]	[0.249]	[0.441]	[0.442]	[0.441]
(log of) Patents (90^{th} percentile)	−0.220**	−0.210*	−0.194	−0.169	−0.162	−0.132	−0.328	−0.313	−0.523**
	[0.109]	[0.110]	[0.128]	[0.110]	[0.112]	[0.131]	[0.204]	[0.203]	[0.209]
R&D incentive (0-3 years)			33.998*			25.927			60.460**
			[17.850]			[18.744]			[24.012]
R&D user cost (0-3 years)			−11.248**			−10.935*			−21.681**
			[5.508]			[5.682]			[8.238]
High-tech incentive × R&D user cost			−9.972*			−7.631			−17.582**
			[5.211]			[5.471]			[6.990]
GDP growth rate (net infrastructure)	0.202***	0.203***	0.113**	0.196***	0.196***	0.112**	0.152**	0.158**	0.070
	[0.040]	[0.040]	[0.035]	[0.041]	[0.041]	[0.037]	[0.077]	[0.077]	[0.067]
Hansen J test (p-value)	(0.395)	(0.436)	(0.277)	(0.527)	(0.552)	(0.362)	(0.847)	(0.859)	(0.405)
Underidentification test (p-value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Exogeneity test (p-value)	(0.725)	(0.734)	(0.558)	(0.812)	(0.814)	(0.786)	(0.646)	(0.649)	(0.203)
AR Weak IV test (p-value)	(0.303)	(0.330)	(0.047)	(0.488)	(0.509)	(0.090)	(0.149)	(0.145)	(0.070)
R-squared (uncentered)	0.204	0.206	0.458	0.196	0.196	0.435	0.071	0.073	0.293
Number of states	45	45	45	45	45	45	45	45	44
Number of observations	812	812	812	805	805	805	752	752	739
Geography controls	N	Y	Y	N	Y	Y	N	Y	Y
Socioeconomic controls	N	N	Y	N	N	Y	N	N	Y
Institutional controls	N	N	Y	N	N	Y	N	N	Y
Supply controls	N	N	Y	N	N	Y	N	N	Y

Notes:
a. All models include Census divisions fixed effects. Additionally, since annual public expenditures vary in sample size, I include a dummy break to account for time fixed effects without compromising the number of covariates. Indeed, its inclusion (not reported) was highly significant in almost all specifications, meaning that this effect cannot be disregarded. Geography controls include average temperatures in winter (January) and summer (July) while socioeconomic ones regard the high school and college graduation rates from total state population. Last but not least, we account for the institutional quality by employing all the disaggregated scores (i.e., subcategories) of the EFNA named: Government spending, Taxation and Labor Market Freedom. Finally, we control for supply shocks with a dummy variable equal to one when the economy had downturns periods (e.g., 1990-91, 2001 and 2008-2010) and zero in contrary case.
b. Real capital outlays of highway spending includes: maintenance, operation, purchases of equipment, toll highways, bridges, tunnels, ferries, street lighting, snow and ice removal. The real measure was computed using the State price index provided by the Bureau of Economic Analysis (BEA) (table 3.9.4) and rescaled into logarithms. Real private non-residential spending comes from U.S Census and deflated by the BLS consumer price index CPI-U (2012=100) and rescaled by state population to control for business cycle effects. Data starts from 1993 onwards.
c. Job-creation rates (private employment) were calculated as the difference of the ratio in employment levels in period t with respect to t-1 times 100. Source: Quarterly Census of Employment & Wages (QCEW).
d. We concentrate on the most innovative high-tech industries (relative to aggregate value share) of patents and scientific citations of all software and telecommunications industries). Thus, our top 4 group, based on the three-digit North American Industry Classification Structure (NAICS) includes the 334, 511, 517 and 541 while the bottom 4 group encompass 512, 515, 518 and 519 sectors.
e. For the high-tech incentive dummy we consider those states in which R&D tax-credits were positive. Data comes from the Panel Database of Incentives and Taxes (PDIT) Bartik (2017). While R&D user cost was constructed following Wilson (2009). For the expenditure share (s) we use COMPUSTAT data (instead of IRS) which is on average 0.1243 between the years 1965 and 2015 rather than 0.5. Particularly, we concentrate on the upper bound (90^{th} percentile) which is about half of the IRS income data (i.e., 0.2588).
f. Autocorrelated standard errors (HAC) with the Newey-West bandwidth adjustment in brackets small sample adjustment. The number of optimal lags for the bandwidth is set to 3.
g. The analysis excludes States with missing data (West Virginia, Wyoming), outliers (Delaware, Maine), the District of Columbia (DC) and States not part of the contiguous United States (Alaska and Hawaii). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: author own calculations.

Table 5: Infrastructure, employment and incentives: IV arbitrary spatial estimates

dv: job creation rate	All			Top value-added			Bottom value-added		
	(I)	(II)	(III)	(I)	(II)	(III)	(I)	(II)	(III)
(log of) Highway spending	−0.192	−0.226	−0.107	−0.210	−0.234	−0.285	−2.339***	−2.368***	−1.579**
	[0.272]	[0.263]	[0.349]	[0.276]	[0.270]	[0.355]	[0.642]	[0.628]	[0.641]
(log of) Private non-residential spending	0.729**	0.720**	0.369	0.917**	0.921**	0.561	0.581	0.448	0.526
	[0.285]	[0.291]	[0.445]	[0.273]	[0.280]	[0.373]	[0.546]	[0.530]	[0.790]
(log of) Policy diffusion score	0.774**	0.784**	0.727**	0.851**	0.853**	0.793**	0.358	0.431	0.047
	[0.304]	[0.302]	[0.260]	[0.310]	[0.309]	[0.270]	[0.569]	[0.572]	[0.522]
R&D incentive (0-3 years)			43.881**			27.184**			85.034**
			[13.247]			[12.432]			[25.720]
R&D user cost (0-3 years)			−1.926			−4.339			−1.495
			[4.461]			[3.432]			[9.862]
R&D incentive × R&D user cost			−12.891**			−8.034**			−24.726**
			[3.847]			[3.620]			[7.460]
GDP growth rate (net infrastructure)	0.178***	0.179***	0.112***	0.178***	0.178***	0.100***	0.166**	0.174**	0.087
	[0.028]	[0.029]	[0.026]	[0.031]	[0.031]	[0.028]	[0.072]	[0.073]	[0.063]
Kleibergen-Paap F statistic	460.600	390.100	223.000	443.400	372.700	210.600	333.700	308.300	192.600
Weak IV two-step test	[−0.408;	[−0.341;	[−0.557;	[−0.414;	[−0.350;	[−0.746;	[−3.215;	[−3.055;	[−2.650;
[2sls intervals]	0.560]	0.667]	0.572]	0.606]	0.668]	0.413]	−1.033]	−0.854]	−0.080]
CSD_W statistic	4.711	0.251	−1.452	0.473	0.605	−2.485	−2.433	−1.798	−0.928
R-squared (uncentered)	0.219	0.219	0.360	0.216	0.216	0.367	0.067	0.070	0.175
Number of states	48	48	48	48	48	48	48	48	48
Number of observations	1007	1007	1007	992	992	992	893	893	893
Geography controls	N	Y	Y	N	Y	Y	N	Y	Y
Socioeconomic controls	N	N	Y	N	N	Y	N	N	Y
Institutional controls	N	N	Y	N	N	Y	N	N	Y
Supply controls	N	N	Y	N	N	Y	N	N	Y

Notes:

a. All models include Census divisions fixed effects. Additionally, since annual public expenditures vary in sample size, I include a dummy break to account for time fixed effects without compromising the number of covariates. Indeed, its inclusion (not reported) was highly significant in almost all specifications, meaning that this effect cannot be disregarded. Geography controls include average temperatures in winter (January) and summer (July) while socioeconomic ones regard the high school and college graduation rates from total state population. Last but not least, we account for the institutional quality by employing all the disaggregated scores (i.e. subcategories) of the EFNA named: Government spending, Taxation and Labor Market Freedom. Finally, we control for supply shocks with a dummy variable equal to one when the economy had downturns periods (e.g., 1990-91, 2001 and 2008-2010) and zero in contrary case.

b. Real capital outlays of highway spending includes: maintenance, operation, purchases of equipment, toll highways, bridges, tunnels, ferries, street lighting, snow and ice removal. The real measure was computed using the State price index provided by the Bureau of Economic Analysis (BEA) (table 3.9.4) and rescaled into logarithms. Real private non-residential spending was obtained from U.S Census and deflated by the BLS consumer price index CPI-U (2012=100) and rescaled by state population to control for business cycle effects. Data starts from 1993 onwards.

c. Job-creation rates (private employment) were calculated as the difference of the ratio in employment levels in period t with respect to t-1 times 100. Source: Quarterly Census of Employment & Wages (QCEW).

d. We concentrate on the most innovative high-tech industries (relative to aggregate value share) of patents and scientific citations of all software and telecommunications industries). Thus, our top 4 group, based on the three-digit North American Industry Classification Structure (NAICS) includes the 334, 511, 517 and 541 while the bottom 4 group encompass 512, 515, 518 and 519 sectors.

e. For the high-tech incentive dummy we consider those states in which R&D tax-credits were positive. Data comes from the Panel Database of Incentives and Taxes (PDIT) (Bartik, 2017). While R&D user cost was constructed following (Wilson, 2009). For the expenditure share (s) we use COMPUSTAT data (instead of IRS) which is on average 0.1243 between the years 1965 and 2015 rather than 0.5. Particularly, we concentrate on the upper bound (90^{th} percentile) which is about half of the IRS income data (i.e. 0.2588).

f. The policy diffusion score includes 204 policies for the following sectors: education, energy, environment, domestic commerce, housing, labor and macroeconomics (Boehmke et al., 2020).

g. To check for the validity of our instruments we employ the LC test for inefficient matrix with a coverage distortion of 5% (i.e. 95% confidence level) and 250 grid points to form a two-step identification-robust confidence set. Results indicate that highway spending estimates always lie within intervals. Thus, the selected instruments show a good performance and statistical inferences can be made.

h. The total number of draws of the weighted CSD statistic were set to 10 to avoid distortions from lower order terms. In this way, the statistic can be constructed accounting for both serial correlation and CSD (Juodis and Reese, 2021).

i. The District of Columbia (DC) as well as Alaska and Hawaii (not part of the contiguous United States) were excluded from the analysis. Standard errors HAC corrected for arbitrary cluster correlation in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: author own calculations.

Table 6: Infrastructure, employment, incentives and innovation: IV arbitrary spatial estimates

dv: job creation rate	All			Top value-added			Bottom value-added		
	(I)	(II)	(III)	(I)	(II)	(III)	(I)	(II)	(III)
(log of) Real highway spending	-0.906^{**}	-0.921^{**}	-0.736^{*}	-0.818^{**}	-0.823^{**}	-0.744^{*}	-2.768^{**}	-2.791^{**}	-2.034^{**}
	[0.363]	[0.368]	[0.430]	[0.371]	[0.374]	[0.428]	[0.825]	[0.833]	[0.876]
(log of) Real non-residential spending	1.024^{**}	0.960^{**}	1.024^{**}	1.156^{**}	1.107^{**}	0.941^{**}	0.548	0.413	1.276
	[0.349]	[0.347]	[0.446]	[0.366]	[0.368]	[0.474]	[0.587]	[0.577]	[0.806]
(log of) Patents (10^{th} percentile)	0.022	0.008	0.037	−0.016	−0.024	0.003	0.221	0.204	0.186
	[0.081]	[0.083]	[0.069]	[0.075]	[0.077]	[0.072]	[0.218]	[0.222]	[0.162]
(log of) Patents (25^{th} percentile)	-0.431^{**}	-0.417^{**}	-0.445^{**}	−0.291	−0.284	−0.303	-1.258^{*}	-1.242^{*}	-1.357^{**}
	[0.190]	[0.192]	[0.188]	[0.193]	[0.195]	[0.212]	[0.660]	[0.660]	[0.666]
(log of) Patents (75^{th} percentile)	0.471^{**}	0.462^{**}	0.485^{**}	0.361^{**}	0.355^{**}	0.322^{**}	0.715	0.674	1.167^{**}
	[0.161]	[0.164]	[0.142]	[0.150]	[0.154]	[0.140]	[0.436]	[0.424]	[0.370]
(log of) Patents (90^{th} percentile)	-0.188^{**}	-0.180^{**}	-0.255^{**}	-0.143^{*}	-0.138^{*}	-0.172^{*}	-0.391^{**}	-0.369^{**}	-0.590^{**}
	[0.076]	[0.076]	[0.099]	[0.076]	[0.076]	[0.098]	[0.188]	[0.185]	[0.188]
R&D incentive (0-3 years)			28.019^{**}			20.841^{*}			54.448^{**}
			[11.195]			[12.485]			[22.617]
R&D user cost (0-3 years)			-11.553^{**}			-9.225^{**}			-22.703^{**}
			[4.277]			[3.998]			[6.641]
R&D incentive × R&D user cost			-8.253^{**}			-6.150^{*}			-15.890^{**}
			[3.261]			[3.635]			[6.532]
GDP growth rate (net infrastructure)	0.208^{***}	0.210^{***}	0.103^{***}	0.202^{***}	0.204^{***}	0.094^{**}	0.200^{**}	0.204^{**}	0.057
	[0.034]	[0.034]	[0.026]	[0.038]	[0.038]	[0.029]	[0.080]	[0.080]	[0.067]
Kleibergen-Paap F statistic	229.800	225.300	161.100	229.000	224.400	152.200	216.500	211.500	157.500
Weak IV two-step test	[−1.498;	[−1.573;	[−1.602;	[−1.315;	[−1.468	[−1.586;	[−4.541;	[−4.450;	[−3.781;
[2sls intervals]	−0.081]	−0.087]	−0.160]	0.170]	; 0.099]	−0.118]	−1.586]	−1.395]	−0.712]
R-squared (uncentered)	0.207	0.208	0.391	0.198	0.199	0.372	0.086	0.088	0.251
Number of states	48	48	48	48	48	48	48	48	48
Number of observations	833	833	833	826	826	826	764	764	764
Geography controls	N	Y	Y	N	Y	Y	N	Y	Y
Socioeconomic controls	N	N	Y	N	N	Y	N	N	Y
Institutional controls	N	N	Y	N	N	Y	N	N	Y
Supply controls	N	N	Y	N	N	Y	N	N	Y

Notes:

a. All models include Census divisions fixed effects. Additionally, since annual public expenditures vary in sample size, I include a dummy break to account for time fixed effects without compromising the number of covariates. Indeed, its inclusion (not reported) was highly significant in almost all specifications, meaning that this effect cannot be disregarded. Geography controls include average temperatures in winter (January) and summer (July) while socioeconomic ones regard the high school and college graduation rates from total state population. Last but not least, we account for the institutional quality by employing all the disaggregated scores (i.e. subcategories) of the EFNA named: Government spending, Taxation and Labor Market Freedom. Finally, we control for supply shocks with a dummy variable equal to one when the economy had downturns periods (e.g., 1990-91, 2001 and 2008-2010) and zero in contrary case.

b. Real capital outlays of highway spending includes: maintenance, operation, purchases of equipment, toll highways, bridges, tunnels, ferries, street lighting, snow and ice removal. The real measure was computed using the State price index provided by the Bureau of Economic Analysis (BEA) (table 3.9.4) and rescaled into logarithms. Real private non-residential spending was obtained from U.S Census and deflated by the BLS consumer price index CPI-U (2012=100) and rescaled by state population to control for business cycle effects. Data starts from 1993 onwards.

c. Job-creation rates (private employment) were calculated as the difference of the ratio in employment levels in period t with respect to t-1 times 100. Source: Quarterly Census of Employment & Wages (QCEW).

d. We concentrate on the most innovative high-tech industries (relative to aggregate value share) of patents and scientific citations of all software and telecommunications industries). Thus, our top 4 group, based on the three-digit North American Industry Classification Structure (NAICS) includes the 334, 511, 517 and 541 while the bottom 4 group encompass 512, 515, 518 and 519 sectors.

e. For the high-tech incentive dummy we consider those states in which R&D tax-credits were positive. Data comes from the Panel Database of Incentives and Taxes (PDIT) (Bartik, 2017). While R&D user cost was constructed following (Wilson, 2009). For the expenditure share (s) we use COMPUSTAT data (instead of IRS) which is on average 0.1243 between the years 1965 and 2015 rather than 0.5. Particularly, we concentrate on the upper bound (90^{th} percentile) which is about half of the IRS income data (i.e. 0.2588).

f. To check for the validity of our instruments we employ the LC test for inefficient matrix with a coverage distortion of 5% (i.e. 95% confidence level) and 250 grid points to form a two-step identification-robust confidence set. Results indicate that highway spending estimates always lie within intervals. Thus, the selected instruments show a good performance and statistical inferences can be made.

g. The total number of draws of the weighted CSD statistic were set to 10 to avoid distortions from lower order terms. In this way, the statistic can be constructed accounting for both serial correlation and CSD (Juodis and Reese, 2021).

h. The District of Columbia (DC) as well as Alaska and Hawaii (not part of the contiguous United States) were excluded from the analysis. Standard errors HAC corrected for arbitrary cluster correlation in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: author own calculations.

Table 7: Spatial autocorrelation test

Variables\Period	1990-94	1995-99	2000-04	2005-09	2010-15
	Moran I statistic				
Employment growth (all high-tech)	0.416***	0.049	0.011	0.046	0.019
(log of) Highway spending	0.065	0.061	0.083*	0.125**	0.141**
(log of) Local aid highway	0.135**	0.205**	0.169**	0.113**	0.063
(log of) Real non-residential spending	0.025	0.056*	−0.007	0.038	0.057
(log of) Policy diffusion score	0.154**	0.093*	0.065	−0.004	0.071
R&D user cost	0.068	0.073	0.081*	0.099**	0.085*

Indicate *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: author own calculations.

Table 8: Infrastructure, employment, incentives and innovation: SLX Estimates

Direct effects	(I)	(II)	(III)	(IV)
(log of) Highway spending	0.154*	0.129	0.151*	0.140*
	[0.085]	[0.082]	[0.084]	[0.082]
(log of) Real non-residential spending	0.049	0.051	0.048	0.051
	[0.047]	[0.047]	[0.046]	[0.046]
Patents top value-added (dummy)	0.616**			
	[0.268]			
Patents top bottom-added (dummy)		0.510*		
		[0.295]		
GDP growth rate (net infrastructure)	0.016***	0.016***	0.017***	0.017***
	[0.005]	[0.005]	[0.005]	[0.005]
Indirect effects & Interactions				
W (log of) Highway spending × (log of) Local aid highway	−0.001	−0.002	−0.002	−0.002
	[0.002]	[0.002]	[0.002]	[0.002]
W Incentives (non-monetary) × Patents (top value-added)	0.183**			
	[0.083]			
W Incentives (non-monetary) × Patents (bottom value-added)		0.152		
		[0.096]		
W R&D (monetary incentive)			0.424**	0.431**
			[0.171]	[0.167]
W R&D user cost × Patents (top value-added)			0.026	
			[0.062]	
W R&D user cost × Patents (bottom value-added)				0.033
				[0.043]
Hansen test (p-value)	0.623	0.603	0.661	0.656
Exogeneity test of instruments (p-value)	0.838	0.804	0.836	0.809
Underidentification test (p-value)	(0.000)	(0.000)	(0.000)	(0.000)
Weak IV AR test (p-value)	(0.352)	(0.423)	(0.381)	(0.411)
CSD statistic	−1.470	−1.710	−1.540	−1.520
R-squared (uncentered)	0.489	0.490	0.490	0.491
Number of states	48	48	48	48
Number of observations	864	864	864	864

Notes:

a. I restrict the analysis to all high-tech sectors with non-missing employment data. I applied the BN first difference (standarized) filtering method (Bai and Ng (2002), Bai (2004) and Bai and Ng, 2004) to control for all time effects. The BN (IC2) criteria yields 3 factors. All models include Census divisions fixed effects, geography and aggregate institutional scores as additional controls. Spillover effects are depicted with the W prefix.

b. All spatial lag variables are of order (1), using a power matrix weight (2) and an arbitrary threshold in kilometres. Highway expenditures are instrumented with the same exogenous variables as in non-spatial models. To account for cumulative effects we augment and restrict its number of past predictors only from 4 to 8 lags. Local aid highway are intergovernmental revenues used for roads, streets, and highways (Pierson et al., 2015).

c. Real private non-residential spending was obtained from U.S Census and deflated by the BLS consumer price index CPI-U (2012=100) and rescaled by state population to control for business cycle effects.

d. Patents is a dummy variable equal to 1 for innovations in the respective (conditional) value-share subsectors and 0 for other aggregated industries: Retail & Construction, Healthcare, Manufacturing and Others (transportation, agriculture, finance and mostly services).

e. Non-monetary incentives include the diffusion of 204 policies in education, energy, environment, domestic commerce, housing, labor and macroeconomics at a State level (Boehmke et al., 2020).

f. R&D incentive dummy for high-tech industries is equal to 1 for those states in which tax-credits were positive (Bartik, 2017). R&D user cost was constructed following (Wilson, 2009) and COMPUSTAT 90^{th} percentile expenditure share (s) (1965-2015) which is about half of the IRS income data (i.e. 0.2588) which is on average 0.1243. Robust standard errors in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: author own calculations.

Table 9: Normality tests for panel data models

	Error (innovation term)			State (specific term)		
Panel A: Top value-added	Skewness	Kurtosis	Joint (p-value)	Skewness	Kurtosis	Joint (p-value)
Coefficient	-0.939^{***}	2.337^{***}	140.530^{***}	-2.291^{***}	6.727^{**}	25.440^{***}
Std.Error	[0.114]	[0.274]	(0.000)	[0.528]	[2.622]	(0.000)
Panel B: Bottom value-added	Skewness	Kurtosis	Joint (p-value)	Skewness	Kurtosis	Joint (p-value)
Coefficient	-0.405^{**}	0.328^{*}	15.370^{**}	-1.777^{***}	2.649^{*}	24.150^{***}
Std.Error	[0.117]	[0.177]	(0.001)	[0.384]	[1.598]	(0.000)

Notes:
We perform the analysis for our two industries using the same controls from employment models and dummies by the four U.S aggregated regions (Northeast, Midwest, South and West) and decades. Bootstrap standard errors (500 replications) in brackets while p-values are in parenthesis. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: author own calculations.

Table 10: Quantile wage decomposition: High-tech industries

Panel A: Top value-added	IV fixed effects	Nonadditive fixed effects				
dv: (log of) Real payroll	(Mean)	q(10)	q(20)	q(50)	q(75)	q(95)
(log of) Highway spending	0.025	−0.113***	−0.052***	0.004***	0.002	0.041***
	[0.073]	[0.003]	[0.004]	[0.001]	[0.002]	[0.011]
(log of) Private non-residential spending	0.028	0.031***	0.088***	0.088***	0.127***	0.097***
	[0.057]	[0.005]	[0.001]	[0.002]	[0.007]	[0.028]
(log of) Patents per-capita	0.017	0.053***	0.035***	0.026***	0.010***	−0.011
	[0.013]	[0.001]	[0.000]	[0.000]	[0.001]	[0.007]
(log of) Policy diffusion score	−0.063	0.029***	0.005**	−0.013***	−0.035***	−0.144***
	[0.044]	[0.005]	[0.002]	[0.001]	[0.002]	[0.011]
R&D incentive (0-3 years)	−2.103	2.005***	0.170***	−0.494***	−2.143***	−4.432***
	[1.375]	[0.004]	[0.004]	[0.001]	[0.004]	[0.057]
R&D user cost (0-3 years)	−2.409***	0.609***	−0.218***	−0.661***	−0.023	0.148
	[0.530]	[0.041]	[0.012]	[0.004]	[0.066]	[0.091]
R&D incentive × R&D user cost	0.579	−0.593***	−0.067***	0.124***	0.601***	1.210***
	[0.398]	[0.000]	[0.001]	[0.000]	[0.002]	[0.011]
Observations	810	847	847	847	847	847
Number of states	46	46	46	46	46	46
Panel B: Bottom value-added	IV fixed effects	Nonadditive fixed effects				
dv: (log of) Real payroll	(Mean)	q(10)	q(20)	q(50)	q(75)	q(95)
(log of) Highway spending	−0.120	−0.115***	0.017***	−0.150***	−0.232***	0.189*
	[0.085]	[0.004]	[0.006]	[0.009]	[0.026]	[0.109]
(log of) Private non-residential spending	0.001	−0.255***	0.004	−0.015**	−0.094***	0.197
	[0.118]	[0.011]	[0.006]	[0.007]	[0.016]	[0.171]
(log of) Patents per-capita	0.018	0.041***	0.027***	0.055***	0.051***	0.069***
	[0.029]	[0.001]	[0.001]	[0.003]	[0.005]	[0.013]
(log of) Policy diffusion score	−0.150**	−0.241***	−0.304***	−0.188***	−0.037**	−0.348***
	[0.072]	[0.004]	[0.006]	[0.006]	[0.019]	[0.079]
R&D incentive (0-3 years)	5.088	3.056***	−4.982***	4.572***	4.152***	3.400***
	[4.370]	[0.003]	[0.007]	[0.004]	[0.102]	[1.294]
R&D user cost (0-3 years)	3.490***	2.204***	1.441***	4.665***	6.863***	14.779***
	[1.052]	[0.059]	[0.054]	[0.091]	[0.188]	[2.416]
R&D incentive × R&D user cost	−1.508	−0.878***	1.353***	−1.394***	−0.977***	−2.332***
	[1.274]	[0.002]	[0.003]	[0.002]	[0.020]	[0.310]
Observations	751	788	788	788	788	788
Number of states	46	46	46	46	46	46

Notes:

a. Conditional mean estimates correspond to the two-way (Census divisions and decades) fixed effects model. The Hansen test (not reported) which tests the validity of the instruments is not rejected in any of the two groups.

b. Real capital outlays of highway spending includes: maintenance, operation, purchases of equipment, toll highways, bridges, tunnels, ferries, street lighting, snow and ice removal. The real measure was computed using the State price index provided by the Bureau of Economic Analysis (BEA) (table 3.9.4) and normalized by state population to control for business cycle effects. Real private non-residential spending was obtained from U.S Census and deflated by the BLS consumer price index CPI-U (2012=100) and normalized by state population to control for business cycle effects. Data starts from 1993 onwards.

d. Patents are the weighted sum of aggregate value share (AV) of U.S public firms in 2001. In lieu with economic literature, patents are rescaled per 100,000 population. Data comes from Arora et al. (2017), patents view and a crosswalk of U.S public firms (COMPUSTAT) Dorn et al. (2020).

c. The policy diffusion score includes 204 policies for the following sectors: education, energy, environment, domestic commerce, housing, labor and macroeconomics (Boehmke et al., 2020). To avoid business cycles rolling windows are included.

d. R&D user cost was constructed following Wilson (2009). For the expenditure share (s) we use COMPUSTAT data (instead of IRS) which is on average 0.1243 between the years 1965 and 2015 rather than 0.5. Particularly, we concentrate on the upper bound (90th percentile) which is about half of the IRS income data (i.e. 0.2588).

e. The District of Columbia (DC) as well as Alaska and Hawaii (not part of the contiguous United States) were automatically excluded from the analysis. Robust standard errors in brackets are computed by bootstrap method with 500 replications using the Metropolis-within-Gibbs sampler. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: author own calculations.

Appendix

Non-spatial Models

Table A.1: Infrastructure, employment and incentives: OLS estimates

dv: job creation rate	All			Top value-added			Bottom value-added		
	(I)	(II)	(III)	(I)	(II)	(III)	(I)	(II)	(III)
(log of) Highway spending	-0.235	-0.262	-0.708^{*}	-0.273	-0.297	-0.731^{*}	-2.558^{***}	-2.549^{***}	-1.442^{*}
	[0.231]	[0.220]	[0.374]	[0.240]	[0.231]	[0.409]	[0.600]	[0.596]	[0.749]
(log of) Private non-residential spending	0.681^{**}	0.687^{**}	0.873^{*}	0.874^{**}	0.895^{**}	0.957^{*}	0.566	0.476	0.481
	[0.282]	[0.295]	[0.484]	[0.265]	[0.277]	[0.497]	[0.530]	[0.518]	[0.614]
(log of) Policy diffusion score	0.661^{**}	0.667^{**}	0.590^{*}	0.750^{**}	0.751^{**}	0.603	0.049	0.079	0.853
	[0.305]	[0.303]	[0.326]	[0.313]	[0.311]	[0.364]	[0.606]	[0.614]	[0.800]
R&D incentive (0-3 years)			43.120^{**}			36.663^{*}			69.228^{**}
			[16.628]			[17.687]			[24.699]
R&D user cost (0-3 years)			-5.507			-6.000			-14.125^{**}
			[5.378]			[5.609]			[6.666]
R&D incentive × R&D user cost			-12.621^{**}			-10.749^{*}			-20.168^{**}
			[4.837]			[5.153]			[7.166]
GDP growth rate (net infrastructure)	0.157^{***}	0.156^{***}	0.119^{***}	0.161^{***}	0.159^{***}	0.119^{***}	0.113	0.119	0.061
	[0.026]	[0.027]	[0.023]	[0.028]	[0.029]	[0.026]	[0.082]	[0.083]	[0.057]
R-squared (within)	0.096	0.096	0.346	0.095	0.095	0.316	0.043	0.044	0.213
Wooldridge AR-1 Test (p-value)	(0.001)	(0.001)	(0.016)	(0.005)	(0.004)	(0.047)	(0.062)	(0.053)	(0.048)
Number of states	48	48	48	48	48	48	48	48	46
Number of observations	1055	1053	1053	1038	1036	1036	929	928	902
Geography controls	N	Y	Y	N	Y	Y	N	Y	Y
Socioeconomic controls	N	N	Y	N	N	Y	N	N	Y
Institutional controls	N	N	Y	N	N	Y	N	N	Y
Supply controls	N	N	Y	N	N	Y	N	N	Y

Notes:

a. All models include Census divisions fixed effects. Additionally, since annual public expenditures vary in sample size, I include a dummy break to account for time fixed effects without compromising the number of covariates. Indeed, its inclusion (not reported) was highly significant in almost all specifications, meaning that this effect cannot be disregarded. Geography controls include average temperatures in winter (January) and summer (July) while socioeconomic ones regard the high school and college graduation rates from total state population. Last but not least, we account for the institutional quality by employing all the disaggregated scores (i.e. subcategories) of the EFNA named: Government spending, Taxation and Labor Market Freedom. Finally, we control for supply shocks with a dummy variable equal to one when the economy had downturns periods (e.g., 1990-91, 2001 and 2008-2010) and zero in contrary case.

b. Real capital outlays of highway spending includes: maintenance, operation, purchases of equipment, toll highways, bridges, tunnels, ferries, street lighting, snow and ice removal. The real measure was computed using the State price index provided by the Bureau of Economic Analysis (BEA) (table 3.9.4) and rescaled into logarithms. Real private non-residential spending was obtained from U.S Census and deflated by the BLS consumer price index CPI-U (2012=100) and rescaled by state population to control for business cycle effects. Data starts from 1993 onwards.

c. Job-creation rates (private employment) were calculated as the difference of the ratio in employment levels in period t with respect to t-1 times 100. Source: Quarterly Census of Employment & Wages (QCEW).

d. We concentrate on the most innovative high-tech industries (relative to aggregate value share) of patents and scientific citations of all software and telecommunications industries). Thus, our top 4 group, based on the three-digit North American Industry Classification Structure (NAICS) includes the 334, 511, 517 and 541 while the bottom 4 group encompass 512, 515, 518 and 519 sectors.

e. For the high-tech incentive dummy we consider those states in which R&D tax-credits were positive. Data comes from the Panel Database of Incentives and Taxes (PDIT) (Bartik, 2017). While R&D user cost was constructed following Wilson (2009). For the expenditure share (s) we use COMPUSTAT data (instead of IRS) which is on average 0.1243 between the years 1965 and 2015 rather than 0.5. Particularly, we concentrate on the upper bound (90^{th} percentile) which is about half of the IRS income data (i.e. 0.2588).

f. The policy diffusion score includes 204 policies for the following sectors: education, energy, environment, domestic commerce, housing, labor and macroeconomics. To avoid business cycles rolling windows are included.

g. The District of Columbia (DC) as well as Alaska and Hawaii (not part of the contiguous United States) were excluded from the analysis. Also we account for outliers (Delaware, Maine). Robust standard errors clustered at State level in brackets (Imbens and Kolesar (2016) small sample adjustment). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: author own calculations.

Table A.2: Infrastructure, employment, incentives and innovation: OLS estimates

dv: job creation rate	All			Top value-added			Bottom value-added		
	(I)	(II)	(III)	(I)	(II)	(III)	(I)	(II)	(III)
(log of) Highway spending	−0.641**	−0.650**	−0.903*	−0.635**	−0.636**	−0.899*	−2.244**	−2.257**	−1.516*
	[0.287]	[0.286]	[0.449]	[0.265]	[0.264]	[0.486]	[0.650]	[0.657]	[0.860]
(log of) Private non-residential spending	0.970**	0.926**	1.153*	1.135**	1.108**	1.214*	0.202	0.129	0.544
	[0.339]	[0.341]	[0.566]	[0.343]	[0.345]	[0.608]	[0.546]	[0.543]	[0.543]
(log of) Patents (10^{th} percentile)	0.122	0.025	0.106	−0.004	−0.008	0.099	0.095	0.083	0.187
	[0.075]	[0.076]	[0.090]	[0.069]	[0.071]	[0.097]	[0.205]	[0.210]	[0.108]
(log of) Patents (25^{th} percentile)	−0.521**	−0.499**	−0.404	−0.360	−0.350	−0.370	−0.813	−0.790	−0.759*
	[0.201]	[0.207]	[0.275]	[0.207]	[0.212]	[0.281]	[0.538]	[0.547]	[0.407]
(log of) Patents (75^{th} percentile)	0.468**	0.462**	0.452**	0.348**	0.342**	0.370*	0.768	0.756	1.177**
	[0.156]	[0.162]	[0.187]	[0.145]	[0.151]	[0.186]	[0.513]	[0.508]	[0.409]
(log of) Patents (90^{th} percentile)	−0.203**	−0.197**	−0.107	−0.154*	−0.150*	−0.058	−0.329*	−0.322*	−0.469**
	[0.081]	[0.082]	[0.137]	[0.085]	[0.086]	[0.138]	[0.188]	[0.187]	[0.215]
R&D incentive (0-3 years)			39.587**			34.507*			68.518**
			[17.854]			[19.399]			[23.259]
R&D user cost (0-3 years)			−9.452			−8.734			−19.888**
			[6.387]			[6.731]			[6.334]
R&D incentive × R&D user cost			−11.582**			−10.114*			−19.933**
			[5.175]			[5.631]			[6.764]
GDP growth rate (net infrastructure)	0.175***	0.177***	0.112***	0.173***	0.174***	0.113***	0.093	0.097	0.054
	[0.033]	[0.034]	[0.025]	[0.036]	[0.036]	[0.027]	[0.088]	[0.089]	[0.063]
R-squared (within)	0.090	0.091	0.367	0.081	0.081	0.344	0.036	0.037	0.252
Wooldridge AR-1 Test (p-value)	(0.005)	(0.004)	(0.071)	(0.015)	(0.013)	(0.127)	(0.067)	(0.058)	(0.113)
Number of states	45	45	45	45	45	45	45	45	44
Number of observations	849	848	848	841	840	840	784	783	770
Geography controls	N	Y	Y	N	Y	Y	N	Y	Y
Socioeconomic controls	N	N	Y	N	N	Y	N	N	Y
Institutional controls	N	N	Y	N	N	Y	N	N	Y
Supply controls	N	N	Y	N	N	Y	N	N	Y

Notes:

a. All models include Census divisions fixed effects. Additionally, since annual public expenditures vary in sample size, I include a dummy break to account for time fixed effects without compromising the number of covariates. Indeed, its inclusion (not reported) was highly significant in almost all specifications, meaning that this effect cannot be disregarded. Geography controls include average temperatures in winter (January) and summer (July) while socioeconomic ones regard the high school and college graduation rates from total state population. Last but not least, we account for the institutional quality by employing all the disaggregated scores (i.e. subcategories) of the EFNA named: Government spending, Taxation and Labor Market Freedom. Finally, we control for supply shocks with a dummy variable equal to one when the economy had downturns periods (e.g., 1990-91, 2001 and 2008-2010) and zero in contrary case.

b. Real capital outlays of highway spending includes: maintenance, operation, purchases of equipment, toll highways, bridges, tunnels, ferries, street lighting, snow and ice removal. The real measure was computed using the State price index provided by the Bureau of Economic Analysis (BEA) (table 3.9.4) and rescaled into logarithms. Real private non-residential spending was obtained from U.S Census and deflated by the BLS consumer price index CPI-U (2012=100) and rescaled by state population to control for business cycle effects. Data starts from 1993 onwards.

c. Job-creation rates (private employment) were calculated as the difference of the ratio in employment levels in period t with respect to t-1 times 100. Source: Quarterly Census of Employment & Wages (QCEW).

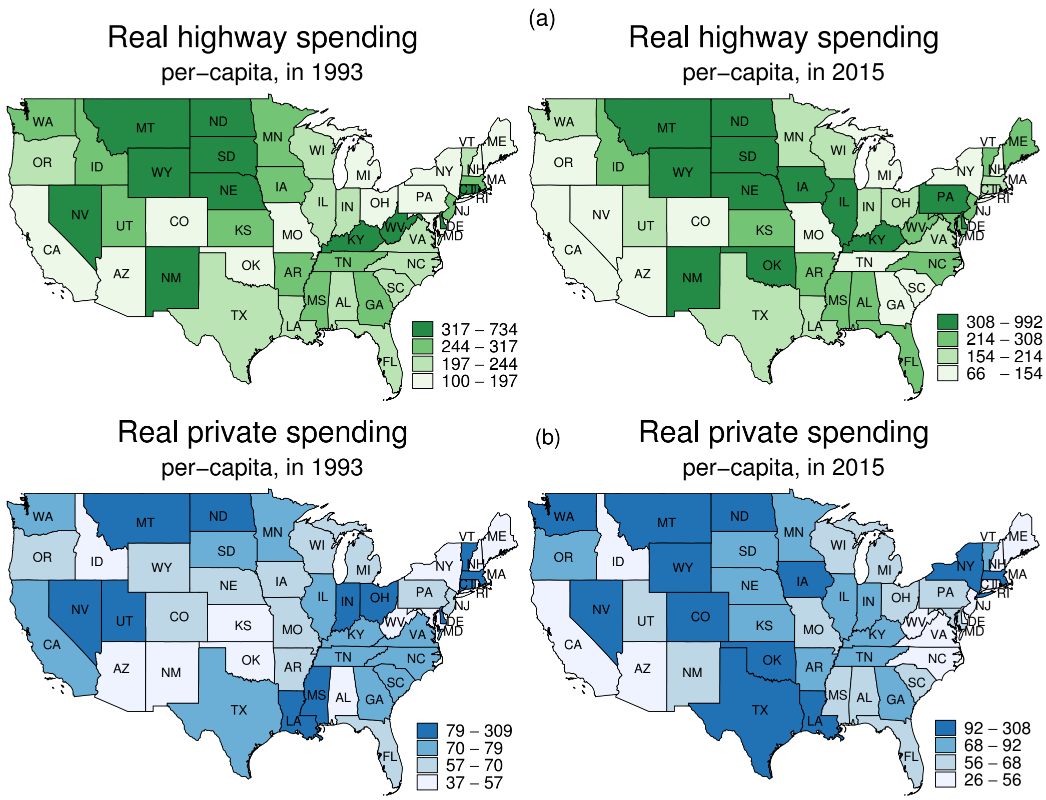
d. We concentrate on the most innovative industries (i.e the distribution —in terms of aggregate value share— of patents and scientific citations of all software and telecommunications industries). Thus, our top 4 group, based on the three-digit North American Industry Classification Structure (NAICS) includes the following sectors: 334, 511, 517 and 541. Conversely, the bottom 4 group encompass 512, 515, 518 and 519.

e. For the high-tech incentive dummy we consider those states in which R&D tax-credits were positive. Data comes from the Panel Database of Incentives and Taxes (PDIT) (Bartik, 2017). While R&D user cost was constructed following Wilson (2009). For the expenditure share (s) we use COMPUSTAT data (instead of IRS) which is on average 0.1243 between the years 1965 and 2015 rather than 0.5. Particularly, we concentrate on the upper bound (90^{th} percentile) which is about half of the IRS income data (i.e. 0.2588).

f. The analysis excludes States with missing data (West Virginia, Wyoming), outliers (Delaware, Maine), the District of Columbia (DC) and States not part of the contiguous United States (Alaska and Hawaii). Robust standard errors clustered at State level in brackets (Imbens and Kolesar (2016) small sample adjustment). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Source: author own calculations.

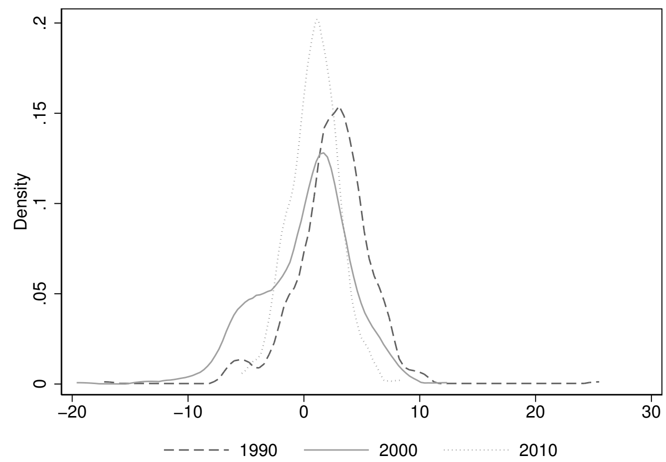
Additional Figures

A.1: Public infrastructure and Private infrastructure



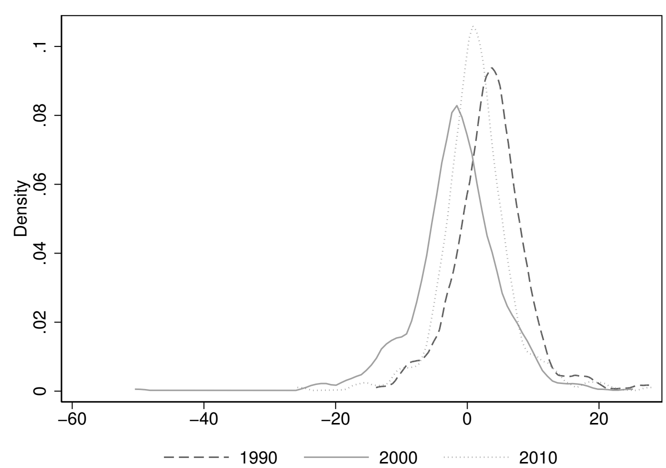
Note: Public infrastructure is defined as highway spending (capital outlays of highway spending which include the following items: maintenance, operation, purchases of equipment, toll highways, bridges, tunnels, ferries, street lighting, snow and ice removal) while private infrastructure is defined as non-residential spending. In the first case, real measures were computed using the State price index provided by the Bureau of Economic Analysis (BEA) which takes into account the price of investment goods (table 3.9.4). While in the second case, the variable was deflated using the BLS consumer price index CPI-U (2012=100). In order to control for business cycle effects both variables are defined in per capita terms. The analysis excludes the District of Columbia (DC) and States not part of the contiguous United States (Alaska and Hawaii). Source: author own calculations.

A.2: Densities of high-tech employment: top value-added industries



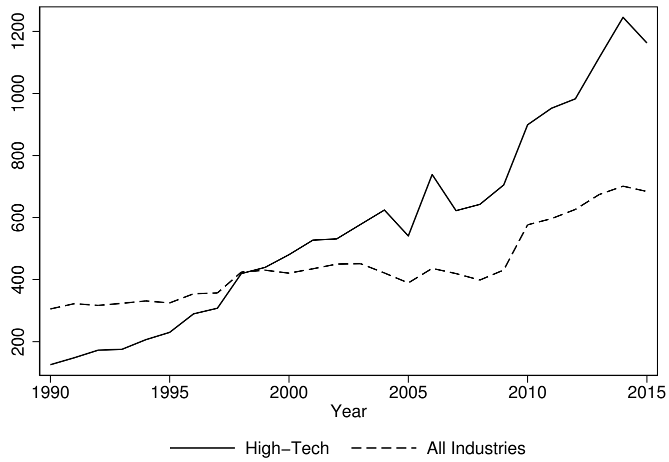
Note: This figure presents kernel density functions for high-tech employment growth rates by decades. Data is from Quarterly Census of Employment & Wages (QCEW). High-tech industries (software and telecommunications) are collapsed at three-digit North American Industry Classification Structure (NAICS) and KLEMS level: 334, 511, 517 and 541 respectively. The analysis excludes the District of Columbia (DC) and States not part of the contiguous United States (Alaska and Hawaii). Source: author own calculations.

A.3: Densities of high-tech employment: bottom value-added industries



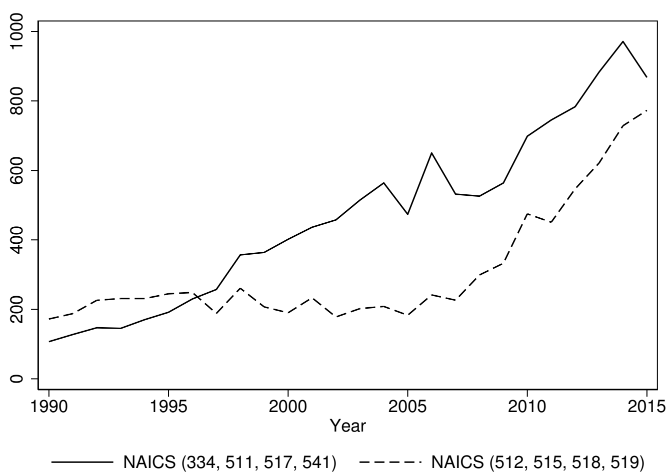
Note: This figure presents kernel density functions for high-tech employment growth rates by decades. Data is from Quarterly Census of Employment & Wages (QCEW). High-tech industries (software and telecommunications) are collapsed at three-digit North American Industry Classification Structure (NAICS) and KLEMS level: 512, 515, 518 and 519 respectively. The analysis excludes the District of Columbia (DC) and States not part of the contiguous United States (Alaska and Hawaii). Source: author own calculations.

A.4: Average number of patents and citations: 1990-2015



Note: Patents data comes from patents view, a crosswalk of U.S public firms (COMPUSTAT) Dorn et al. (2020) and Arora et al. (2017). High-tech industries (software and telecommunications) were collapsed at three-digit North American Industry Classification Structure (NAICS) and KLEMS level: 334 (Computer & electronic product manufacturing); 511 (Publishing industries except internet); 512 (Motion picture & sound recording industries); 515 (Broadcasting except internet); 517 (Telecommunications); 518 (Data processing, hosting, & related services); 519 (Other information services); 541 (Professional, scientific, & technical services). All industries encompass: retail & construction, healthcare, manufacturing (including mining and utilities), and others: finance, transportation warehousing, agriculture and others (mostly services).

A.5: Average number of patents and citations by value added: 1990-2015



Note: Patents data comes from patents view, a crosswalk of U.S public firms (COMPUSTAT) Dorn et al. (2020) and Arora et al. (2017). Data comes from COMPUSTAT (public firms). High-tech industries (software and telecommunications) at three-digit North American Industry Classification Structure (NAICS) level: 334 (Computer & electronic product manufacturing); 511 (Publishing industries except internet); 512 (Motion picture & sound recording industries); 515 (Broadcasting except internet); 517 (Telecommunications); 518 (Data processing, hosting, & related services); 519 (Other information services); 541 (Professional, scientific, & technical services).

Infrastructure, innovation and wages: Measures of shape

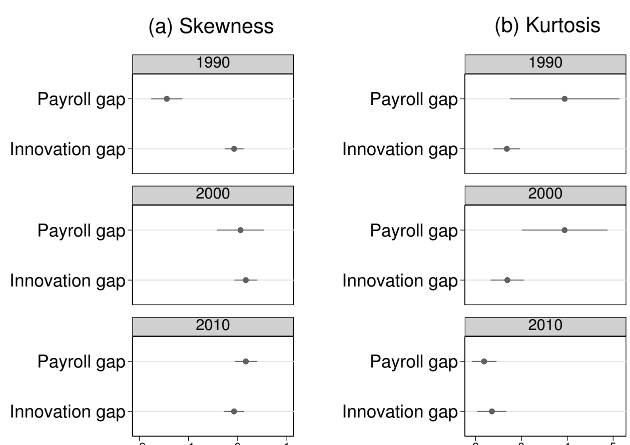
In order to check for normality or its absence, we focus on pairs of relations in terms of gaps, periods (i.e., decades) and regions (i.e., Census divisions). To this end, we first compute the payroll gap as the difference between the 90^{th} and 10^{th} percentile of high-tech real wages and the innovation gap as the difference between 90^{th} and 10^{th} patents and scientific citations in high-tech industries. Unlike the traditional Q-Q plot, here we intent to have a clear picture of dispersion measures between wages and innovation over the period of analysis (1990-2015).

In Figure A.6 panel (a), the dispersion by decades at the beginning of the period (1990) shows a negative skewed distribution on both payroll and innovation gap for the high-tech industry. In other words, a fatter tail on the left side of top earners with respect to bottom ones suggests that top state innovators were slack or not properly compensated in terms of better salaries. However, in the following decades such difference substantially narrowed (2000) to finally rebounded it (2010). The latter situation might indicate that the creation of new knowledge —despite the extensive improvements in technology— has decreasing marginal returns with respect to wages. Conversely, in panel (b) an excess of (positive) kurtosis is shown, but mostly focus on payrolls in the first decades (1990 and 2000) meaning that the presence of outliers is significant. In this line, the ICT revolution is an important driving factor within the high-tech industry. Hence, in our quantile regressions we will have special care when including decadal dummies.

In Figure A.7 we conduct the same exercise but focusing on the dispersion by divisions. In panel (a) results indicate dissimilar skewed distributions across divisions. For example, in West South Central the payroll gap is higher than the innovation gap whereas in the East Central Division the converse is observed. The reason is simple, the share of patents of in those regions are relatively lower (higher) with respect to salaries. Naturally, the latter relation is less clear in the Silicon Valley (i.e., Pacific) for which the number of high-tech firms and the quantity of scientific patents are considerably much higher. Along these lines, panel (b) reinforces the same trend from above as an excess of (positive) kurtosis in regards to innovations is observed. Certainly, the presence of star scientists in some states (e.g., California) and/or the lack of scientific contributions in some others (e.g., Wyoming) are factors which can perfectly explain the non-normality of the data. To this end, quantile panel models with nonadditive fixed effects are perfectly capable to deal with those issues.

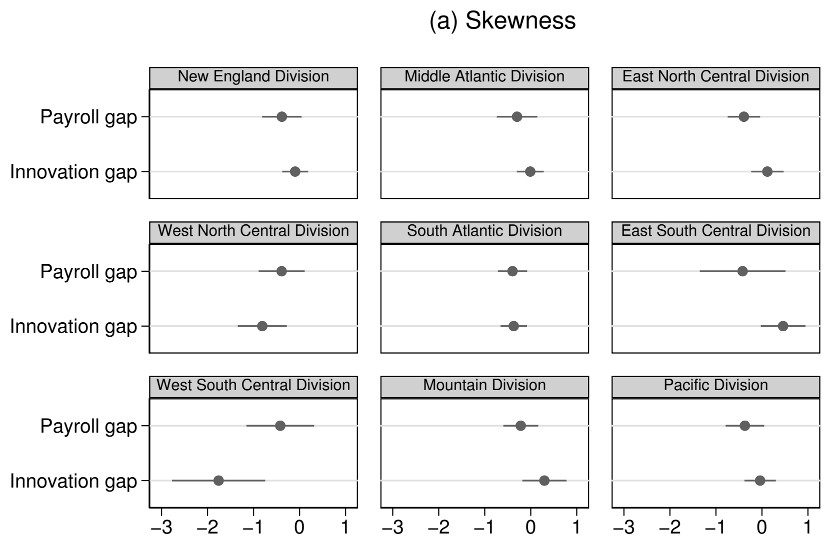
Repeating the same analysis from above, but focusing on payrolls and infrastructure, the story is different. Panel (a) of Figure A.8 shows that the payroll gap in high-tech industries have a slightly negative skewed distribution at the beginning of the sample (1990) while both measures of infrastructure have fatter right tails. However, during the last decade (2010) such difference narrows. In panel (b) of the same figure, again a (positive) excess of kurtosis is displayed. But in this case, high-tech wages during the 2000 are excessively higher than both measures of infrastructure. A possible explanation is the presence of outliers cope with delayed effects of past public policies like R&D tax incentives and/or innovation policies. Turning our attention to divisions, panel (a) from Figure A.9 indicates that highway expenditures tend to be absorbed disproportionally more in bigger states (e.g., New England). Although, the influence of private infrastructure investments has increased over time. In this vein, panel (b) confirms a leptokurtic distribution of private infrastructure investments in three out of six divisions (New England, West North and West South Central). Therefore, even though private investments cannot be directly compared to state traditional ones (e.g., a new highway), the formers should not be disregarded from the analysis as they can also support and facilitate diverse activities performed by firms; thereby providing cost-saving effects and foremost better working conditions to worker's.

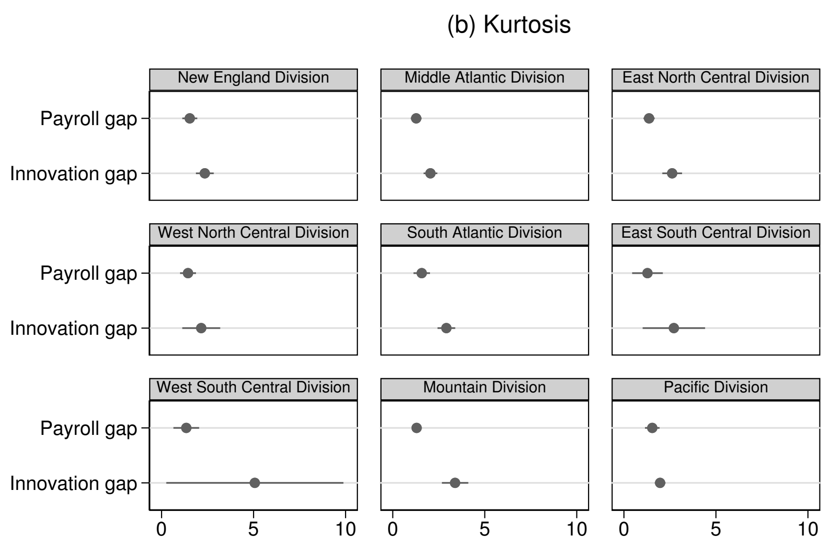
A.6: Wage gap and innovation gap: dispersion by decades



Note: Payroll gap is the difference between the 90^{th} and 10^{th} percentile of real wages; innovation gap is the difference between 90^{th} and 10^{th} patents and scientific citations. Source: Author own calculations based on County Business Paterns (CBP), patents view, a crosswalk of U.S public firms (COMPUSTAT) Dorn et al. (2020) and Arora et al. (2017).

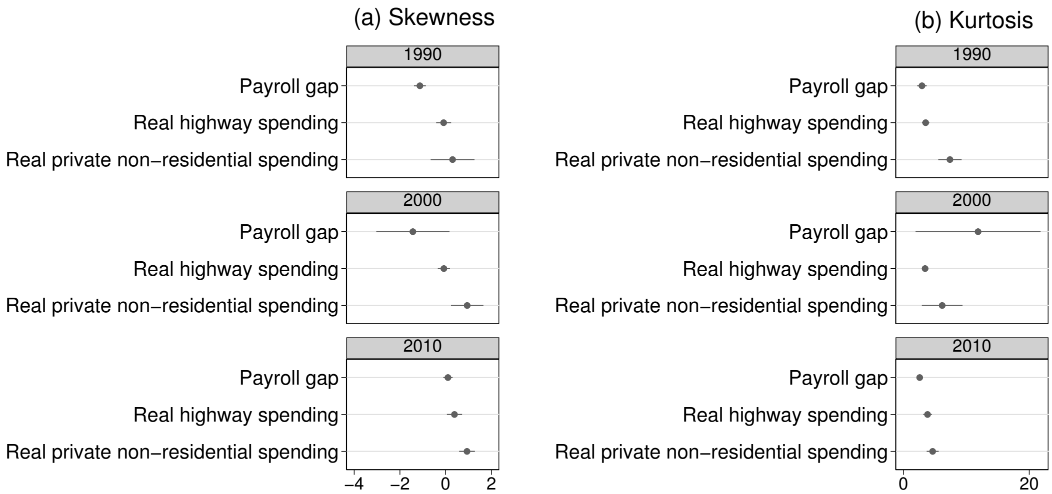
A.7: Wage gap and innovation gap: dispersion by divisions





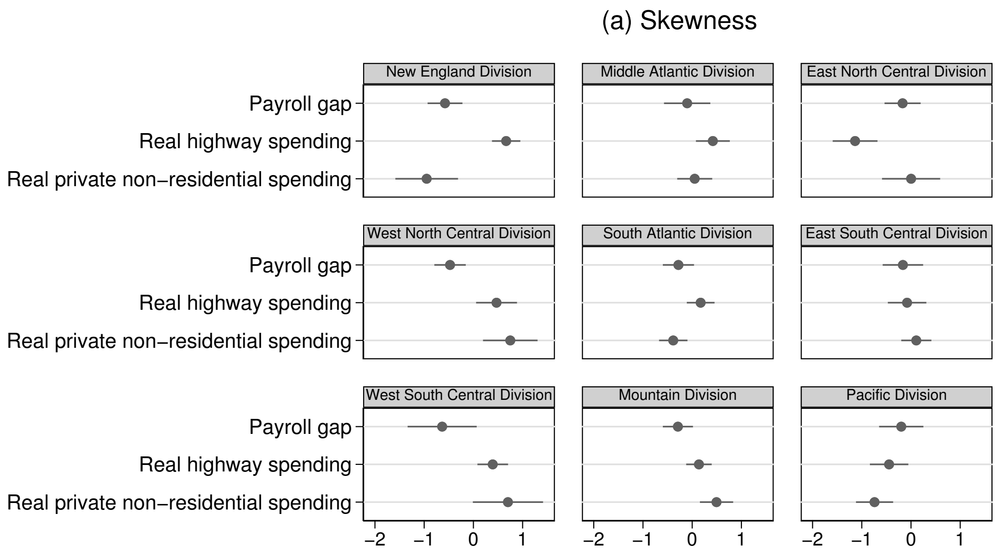
Note: Payroll gap is the difference between the 90^{th} and 10^{th} percentile of real wages; innovation gap is the difference between 90^{th} and 10^{th} patents and scientific citations. Source: Author own calculations based on County Business Paterns (CBP), patents view, a crosswalk of U.S public firms (COMPUSTAT) Dorn et al. (2020) and Arora et al. (2017).

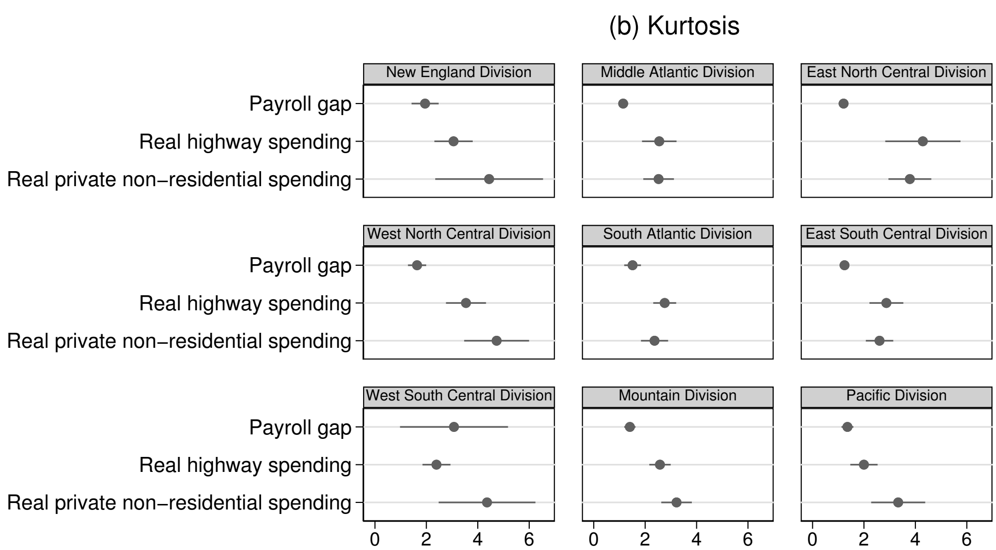
A.8: Wage gap and infrastructure: dispersion by decades



Note: Payroll gap is the difference between the 90^{th} and 10^{th} percentile of real wages; real highway spending includes all capital outlays. Real measures were computed using the State price index (BEA) and the CPI-U (2012=100). Source: Author own calculations based on County Business Paterns (CBP), Pierson et al. (2015) and BLS.

A.9: Wage gap and infrastructure: dispersion by divisions





Note: Payroll gap is the difference between the 90^{th} and 10^{th} percentile of real wages; real highway spending include all capital outlays. Real measures were computed using the State price index (BEA) and the CPI-U (2012=100). Source: Author own calculations based on County Business Paterns (CBP), Pierson et al. (2015) and BLS.