The long-run effects of R&D place-based policies: Evidence from Russian Science Cities*

Helena Schweiger,† Alexander Stepanov‡ and Paolo Zacchia§

April 2020

Abstract

We study the long-run effects of historical place-based policies targeting R&D: the creation of *Science Cities* in former Soviet Russia. The establishment of Science Cities and the criteria for selecting their location were largely guided by idiosyncratic considerations of military and strategic nature. We compare current demographic and economic characteristics of Science Cities with those of appropriately matched localities that were similar to them at the time of their establishment, and had similar pre-trends. We find that in present-day Russia, despite the massive cuts in government support to R&D that followed the dissolution of the USSR, Science Cities host more highly skilled workers and more developed R&D and ICT sectors; they are the origin of more international patents; and they generally appear to be more productive and economically developed. Within a spatial equilibrium framework, we interpret these findings as the result of the interaction between persistence and agglomeration forces. Furthermore, we rule out alternative explanations related to the differential use of public resources, and we find limited evidence of reversion to the mean. Lastly, an analysis of firm-level data suggests that locating closer to Science Cities generates localized spillover effects on firms' innovation and performance indicators.

*We would like to thank Charles Becker, Ekaterina Borisova, Vincenzo Bove, Julian Cooper, Ralph de Haas, Sergei Guriev, Maria Gorban, Mark Harrison, Denis Ivanov, Sergei Izmalkov, Patrick Kline, Olga Kuzmina, Andrei Markevich, Enrico Moretti, Gérard Roland, Mauro Sylos Labini and Natalya Volchkova for helpful discussions, as well as the participants at the SITE Academic Conference: 25 years of transition, 8th MEIDE conference, 2015 PacDev conference, 17th IEA World Congress, 16th Uddevalla Symposium, 2nd World Congress of Comparative Economics, XVIII HSE April International Academic Conference, 2017 Barcelona Workshop on Regional and Urban Economics, 2018 Meeting of the American Economic Association, 13th Meeting of the Urban Economics Association and seminars at the EBRD, Higher School of Economics, New Economic School, U.C. Berkeley, U. of Genoa, U. of Cergy-Pontoise, École Polytechnique, U. of Nottingham, U. of Warwick and CERGE-EI for their comments and suggestions. Irina Capita, Jan Lukšič and Maria Vasilenko provided excellent research assistance. The views expressed in this paper are our own and do not necessarily represent those of the institutions of affiliation.

†European Bank for Reconstruction and Development (EBRD). E-mail: `schweigh@ebrd.com`.

‡European Bank for Reconstruction and Development (EBRD). E-mail: `stepanoa@ebrd.com`.

§CERGE-EI and IMT School of Advanced Studies Lucca. E-mail: `Paolo.Zacchia@cerge-ei.cz`.

JEL Classification Codes: N94, O38, P25, R58

Keywords: Science Cities, Place-based policies, Agglomeration, Russia

1 Introduction

The effectiveness of public support for science and research and development (R&D) is a long-standing issue in the economics of innovation. Both direct subsidies and indirect incentives for research and science are typically predicated upon positive externalities (or other types of market failures) which, in the absence of public intervention, cause under-investment in R&D. Some specific innovation policies, such as the top-down creation of local R&D clusters, are characterized by a geographical local dimension. In such contexts, assessing the spatial extent of knowledge spillovers – one of the three forces of spatial agglomeration first identified by Marshall (1890), corresponding with the "learning" effect from the more recent classification by Duranton and Puga (2004) – is relevant for evaluating the overall effect of the intervention. The debate about localized innovation policies mixes with the one about broader place-based policies. In particular, the focus is on whether place-based policies can succeed at generating self-reinforcing economic effects that persist after their termination, possibly because of agglomeration forces at work. In the absence of long-run effects, the net welfare effect of place-based policies is as likely to be negative as it is to be positive (Glaeser and Gottlieb, 2008).[1]

Localized innovation policies are a popular policy intervention, but evidence about their effectiveness is scarce even in the short-run, let alone the long-run. This paper is one of the few analyzing such policies and their long-run impact. We are able to do so by examining the establishment of highly specialized Science Cities in the territory of modern Russia during Soviet times. We identify 95 middle-sized urban centers that were created or developed by the Soviet government with the purpose of concentrating strategic R&D facilities. Each city was typically shaped around a specific technological purpose; in order to work in the newly created establishments, the Soviet government relocated scientists, researchers and other high-skilled workers from elsewhere in the Soviet Union. The creation of Science Cities was motivated by the technological and

[1] Their argument is based on the interaction between congestion effects and spatial agglomeration externalities – such as those due to local knowledge spillovers – in a spatial equilibrium model that allows for movement of workers across places. In their theoretical framework, place-based policies are interpreted as a reallocation of employment between areas, hence they are welfare-improving only if the benefits accrued to the target regions are larger than the costs experienced elsewhere. This, in turn, is possible as long as agglomeration economies more than countervail the congestion effects as employment increases. The non-linearities implicit in this condition entail multiple equilibria, thus, place-based policies can be seen as "equilibrium shifters". This has motivated subsequent empirical research aimed at uncovering agglomeration effects and their (potential) non-linearities. See also the discussion in Glaeser and Gottlieb (2009) and Kline and Moretti (2014b).

military competition between competing geopolitical blocks in the context of the Cold War; it is thus unsurprising that most of them were specialized in military-related fields, such as nuclear physics, aerospace, ballistics and chemistry. These sectors remain, to this day, those in which Russia maintains a comparative technological advantage.

While one may question whether the institutional context of Russian Science Cities is comparable to that of other industrialized countries, this historical experience actually works in favor of estimating the long-run causal impact of a place-based policy. First, it greatly diminishes concerns for selection biases due to unobserved determinants of future development, which typically affect studies about innovative clusters in other countries. An analysis by Gregory and Harrison (2005) reveals that the allocation of resources in the Soviet command economy was managed according to suboptimal, often erratic rules of thumb – especially so for highly secretive projects, which were managed by a handful of bureaucrats lacking the advice of experts (Harrison, 2018). Soviet science was, in fact, a largely secretive business managed by the internal police and the army (Siddiqi, 2008, 2015). A rule of thumb appears most evident in the choice of Science Cities' locations (Agirrechu, 2009), which was based on a secrecy-usability trade-off: the Soviet leaders prioritized places that offered better secrecy and safety from foreign interference (in the form of R&D espionage), or that were otherwise easy to control by governmental agencies, by virtue of geographical proximity. The potential for economic development and local human capital accumulation was arguably not, at the margin, a determinant of a location's choice for the establishment of a Science City.

Second, the transition to a market economy that followed the dissolution of the USSR resulted in a large negative shock for Russian R&D, as direct governmental expenditure in R&D as a percentage of GDP fell by about 75 per cent, causing half of the scientists and researchers of post-1991 Russia to lose their jobs. Consequently, state support for Science Cities was abruptly suspended for obviously exogenous reasons; only in the 2000s was it partially resumed for 14 of the former towns, which today bear the official name of Naukogrady (Russian for "Science Cities"). These historical developments provide us with a unique opportunity to study the long-run consequences of an exogenous spatial reallocation of highly-skilled workers, decades after the termination of the program that originally motivated such reallocation. In addition, by analyzing historical Science Cities separately from modern Naukogrady, we are able to evaluate to what extent the modern characteristics of the former depend on the long-run effects due to the Soviet-era policy, rather than on current government support.

These distinctive institutional features motivated us to build a unique, rich dataset covering geographic, historical and present characteristics of Russian municipalities in order to answer more general questions about the consequences of innovation-centered place-based policies. Specifically, our contribution to the extant literature is three-fold. First, we estimate the long-run impact of the past establishment of Science Cities on various demographic and economic characteristics of the selected municipalities, finding largely positive effects measured twenty years after the transition. Second, we test other channels that might explain our findings, including the possible role of a differential use of public resources – a hypothesis put forward by von Ehrlich and Seidel (2018, henceforth 'vES'). We largely rule out these alternative mechanisms. Third, we examine spillovers of R&D place-based policies on the productivity and innovation of businesses that operate in the market economy of today's Russia, obtaining correlational evidence that Science Cities do positively affect firms, but within a narrow distance in space.

We estimate the effect of the past establishment of a Science City on present-day municipal-level human capital (measured as the share of the population with either graduate or postgraduate qualifications), innovation (evaluated in terms of patent output measures) and various proxies of economic development. In order to give a causal interpretation to our estimates, we construct an appropriate control group by employing matching techniques. In particular, we match Science Cities to other localities that, at the time of selection, were similar to them in terms of characteristics that could affect both their probability of being chosen and their future outcomes, and were on a similar trend of population growth. Our main identifying assumption is that, conditional on our matching variables, the choice of a locality was determined at the margin by factors that would be independent from future, post-transition outcomes. In order to implement this strategy, we construct a unique dataset of Russian municipalities, which combines geographical, historical and more recently observed local characteristics.

Our results can be summarized as follows. At present, Science Cities from the Soviet era still host a more educated population, are more economically developed, employ a larger number of workers in R&D and ICT-related jobs, and produce more patents than other localities that were comparable to Science Cities at the time of the program's inception. In addition, researchers working in former Science Cities appear to be more productive, and to receive substantially higher salaries. The estimated treatment effect is typically lower than the raw sample difference for all our outcome variables except those related to patents, for which no ex-ante bias can be attested from our estimates.

When we exclude modern Naukogrady from the analysis, the results remain largely unchanged, but the point estimates relative to total and per capita patent production decrease by about 65 per cent. We also perform a more in-depth analysis of demographic outcomes and economic development (proxied by night lights), which reveals little to no evidence of mean reversion.

We interpret the results in light of a spatial equilibrium model à la Glaeser and Gottlieb (2009) and Moretti (2011). In the model, the Soviet Union initially allocates workers of different skills in Science Cities and other localities; after the transition, workers are allowed to move. The model is flexible enough to allow for multiple agglomeration and dispersion forces. The equilibrium predictions of the model are well-suited to simple estimation based on the matched sample; under the maintained identification strategy, this allows us to disentangle and quantify the various mechanisms. The results obtained from this exercise point to a combination of agglomeration and "persistence forces" that induce demographic path-dependence as the explanation of our main findings: following the dissolution of the USSR, cities endowed with a higher concentration of skilled workers – typically unwilling or unable to move – were better suited to the transition to a market economy driven by services. Notably, the related agglomeration elasticities in high-skilled sectors are estimated around 0.08 or higher, in line with the existing literature about middle income countries (Melo et al., 2009; Duranton, 2015).

This interpretation contrasts with the study by von Ehrlich and Seidel (2018) of the formerly subsidised West German municipalities which used to border the Iron Curtain. Their empirical analysis rules out agglomeration effects; instead, they propose persistence in public goods investment as the explanation of their measured long-run effects. Our paper is the first in the literature to provide an assessment of the vES hypothesis by analyzing municipal budget data. We find that, with equal available resources, at least some former Science Cities seem to spend more per capita – with respect to matched towns – on education and school maintenance. Unfavorably for the vES hypothesis, though, these differences are not statistically robust, and no analogous difference can be attested for spending in utilities and physical infrastructure. On the other hand, local expenditures in education – unlike spending in physical infrastructure – appear to be a strong predictor of individual wages when estimating our spatial equilibrium model.

Furthermore, we complement the municipal-level empirical analysis with an additional set of estimates based on firm-level data, to evaluate whether in present-day Russia, the effect of Science Cities spills over onto other firms that are located nearby, and to what economic and geographical extent. We use data on Russian firms from the fifth round of the Business Environment and Enterprise Performance Survey (BEEPS V), which were sampled from the regions where the majority of Science Cities are located. We evaluate to what extent the distance of a firm from a Science City correlates with its innovation and performance outcomes and find that many innovation and performance indicators are enhanced for those firms that are based close to Science Cities. These results reinforce our conclusion that the municipal-level differentials are at least in part caused by knowledge spillovers.

Our paper contributes to various strands of literature. First, we add to the growing number of studies about the evaluation of place-based policies; for a recent survey of the empirical research see Neumark and Simpson (2014).[2] Our paper is most directly related, conceptually and methodologically, to the studies by Kline and Moretti (2014a) on the Tennessee Valley Authority, Fan and Zou (2017) on China's "Third Front" state-driven industrialization of inner China, and Heblich et al. (2018) about the "Million Roubles Plants" built in China with Soviet support during the early Cold War. While these contributions uncover long-run effects from historical place-based policies focusing on physical capital, the Science Cities program stands out as it specifically concerned investments in knowledge. Similarly to these papers, our empirical strategy also exploits unique historical circumstances of political, geographical and military kind in order to construct an appropriate control group for Science Cities.

Second, and relatedly, we contribute to the more general search of agglomeration effects – and in particular of the third Marshallian force, localized knowledge spillovers – in urban and regional economics. This has long been a traditional field of investigation for economic geographers, with a particular interest in innovation clusters. Following seminal contributions by Jaffe (1989), Glaeser et al. (1992), Audretsch and Feldman (1996) and others, a large literature has developed.[3] Recently, the issue has caught the attention of economists working in other fields, with several papers focusing on local productivity spillovers; for example: Moretti (2004), Ellison et al. (2010), Greenstone et al. (2010), Bloom et al. (2013) and Lychagin et al. (2016).

[2]The majority of these papers analyze policies enacted in the United States (Neumark and Kolko, 2010; Busso et al., 2013; Kline and Moretti, 2014a) or in the European Union (Bronzini and de Blasio, 2006; Criscuolo et al., 2012; Givord et al., 2013; von Ehrlich and Seidel, 2018).

[3]We propose two fairly recent surveys: Beaudry and Schiffauerova (2009) focus on the "Marshall vs. Jacobs" debate around the prevalence of, respectively, within- versus between-industry local knowledge spillovers; while Boschma and Frenken (2011) devote special attention to studies within the evolutionary economic geography research agenda.

Third, the institutional setting of this paper relates it to other contributions on the consequences of historically massive forms of government intervention on long-run economic and technological development, be it in Russia or elsewhere. Cheremukhin et al. (2017) argue that the "Big Push" industrialisation policy enacted in the USSR under Stalin did not succeed in shifting Russia onto a faster path of economic development. Mikhailova (2012a,b) evaluates negative welfare effects due to the regional demographic policies enacted in the Soviet Union; she also finds that locations hosting Gulag camps grew significantly faster than similar places without camps. Things look different in the more specific case of R&D policies. Through an analysis performed at a higher level of geographic aggregation than ours, Ivanov (2016) finds that Russian regions with more R&D personnel before the onset of transition do better today at expanding employment in high-tech sectors. Outside Russia, Moretti et al. (2016) show that increases in government-funded R&D for military purposes have positive net effects on the TFP of OECD countries, despite crowding out private expenditures in R&D.

This paper is organized as follows. Section 2 summarizes the history of Science Cities and the process by which their locations were chosen. Section 3 illustrates the analysis of the long-run effects at the municipal level: it outlines the empirical methodology, describes the data employed and discusses the empirical results. Section 4 introduces the conceptual framework and elaborates on the empirical estimates based on its equilibrium equations. Section 5 focuses on the firm-level analysis, again separating between the methodology, the data and the results. Section 6 concludes.

2 Historical and institutional background

In what follows, we provide a historical overview of Science Cities and their institutional context; in addition, we discuss major factors conducive to the choice of their locations.

2.1 Science Cities of the Soviet Union

The former Soviet Union was in a way a pioneer in public investment in science and in place-based policies focusing on R&D. In the context of the Cold War competition between the United States and the USSR, the Soviet leadership prioritised the allocation of the best resources – including human – to sectors considered vital to the country's national security. Around two-thirds of all Soviet R&D spending was set for military purposes, and almost all of the country's high-technology industry was in sectors directly or indirectly related to defence (Cooper, 2012). Moreover, science was mostly a responsibility of the army (Siddiqi, 2015). Science Cities emerged in this environment. We identify 95 middle-sized urban centers which the Soviet government endowed with a high concentration of research and development facilities, each devoted to a particular scientific and technical specialization.[4] Science Cities began to develop around strategically important (military) research centers from the mid-1930s;[5] however, the majority of them were established after the World War II, especially in the 1950s.

As they specialized in industries with high technological intensity, Science Cities needed access to suitable equipment, machinery, intermediate inputs and qualified personnel. With the objective of co-locating scientific research centers, training institutes and manufacturing facilities, the Soviet government established about two-thirds of Science Cities by "repurposing" existing settlements, while the rest were built from scratch in sparsely populated areas. Researchers, scientists and supporting personnel were relocated to Science Cities in order to contribute to the R&D projects. To incentivize them, the Soviet authorities strove to provide better than standard living conditions in these localities, by making available a wider choice of retail goods, more comfortable apartments as well as more abundant cultural opportunities than elsewhere in the country. Typically, the urban characteristics of Science Cities were better than those of other contemporary settlements, as the former were developed according to the best urban planning criteria at the time (Agirrechu, 2009).

Starting in the 1940s, with the need to protect the secrecy of the nuclear weapons program in the Cold War environment (Rowland, 1996; Cooper, 2012), many Soviet municipalities of military importance were "closed" to external access in order to maintain security and privacy. Non-residents needed explicit permission to travel to closed cities and were subject to document checks and security checkpoints; relocating to a closed city required security clearance by the KGB; foreigners were prohibited from entering

[4]The term "Science City" (Naukograd) was first introduced in 1991 (Ruchnov and Zaitseva, 2011). The former Soviet Union was not a true Science Cities pioneer — the first Science City was established in 1937 in Peenemünde, Germany — but it implemented the idea to a much larger extent.

[5]The model of innovation followed by the Soviet authorities since the early 1930s was the creation of "special-regime enclaves intended to promote innovation" (Cooper, 2012). These enclaves first appeared as secret research and development laboratories (so-called Experimental Design Bureaus or sharashki) in the Soviet Gulag labor camp system (Siddiqi, 2008). The scientists and engineers employed in a sharashka were prisoners picked from various camps and prisons, and assigned to work on scientific and technological problems; they were often responsible for technological breakthroughs, although sharashki operated under a suboptimal scheme of incentives (Gregory and Harrison, 2005).

them at all; and inhabitants had to keep their place of residence secret. Science Cities whose main objective was to develop nuclear weapons, missile technology, aircraft and electronics were closed as well; some of them were located in remote areas situated deep in the Urals and Siberia – out of reach of enemy bombers – and were represented only on classified maps. Note that the sets of "Science Cities" and "closed cities" overlap only partially, a fact that we take into account in our empirical analysis.

Following the dissolution of the USSR, Russia underwent a difficult transformation from a planned to a market economy. The withdrawal of the state from many sectors of the economy dramatically affected R&D as well. In Russia, gross R&D expenditures as a fraction of GDP fell from the 1990 level of about 2 per cent to a mere 0.74 per cent in 1992, a fact made even more dramatic as Russian GDP shrank by about 50 per cent in the early years of the transition. Wages plummeted, and consequently, total employment in R&D fell by about 50 per cent.[6] Over the transition, there was little to no recovery from these initial shocks (see Appendix A for more details). This has inevitably affected Science Cities; while we lack access to detailed information about their government funding in the 1990s, anecdotal evidence speaks of an effective discontinuation of the military research programs that Science Cities were responsible for, at least until the government, starting in the early 2000s, re-established direct support for the 14 modern Naukogrady mentioned in the introduction (though this time without explicit military focus). Our analysis of recent municipal budgets (see section 3.3) confirms that Science Cities today receive, if anything, lower governmental transfers than comparable towns; as we elaborate later, we tentatively attribute this fact to political reasons.

2.2 The choice of Science Cities' locations

Since Science Cities were created in secret and in a staggered fashion, typically for the sake of responding to perceived military-technological threats coming from the west, there exists no detailed, comprehensive account of how their locations were chosen.[7] To

[6] Whereas in Soviet times the wages of scientists were about 10-20 per cent higher than the average, they dropped to about 65 per cent of the average wage in 1992, following the withdrawal of the state from the R&D sector (Saltykov, 1997). Even worse, during the 1990s many scientists did not even receive their salary, or received only a fraction of it (sometimes in kind) over extended periods (Ganguli, 2015). Low remuneration was not the only reason for researchers to leave the R&D sector: with the removal of previous restrictions to individual mobility, scientists were allowed to migrate abroad - though that occurred mostly along ethnic lines.

[7] The Russian Ministry of Defense started working on a series of publications documenting the activities of its Soviet predecessor until 1960. However, the publications available so far only cover the period

shed some light on this and inform our empirical analysis, we discuss the top decision processes in the Soviet command economy, drawing on the historical meta-analysis by Gregory and Harrison (2005) which examines the archives produced by the Soviet state between its inception and 1960. The authors make an observation that is very relatable to Science Cities: in the USSR, the allocation of resources was not informed by criteria of efficiency and optimality; instead, it was based on very imperfect and informal "rules of thumb."[8] This was ultimately caused by a combination of Hayekian informational problems, and a failure of the Soviet model of political economy – at all levels of its bureaucratic apparatus – to credibly commit to a set of contingent, efficient rules for the management of a planned economy. In the case of decisions with a content of secrecy, these problems were exacerbated by a security-usability trade-off (Harrison, 2018). In short, the Soviet leaders commonly faced a dilemma: while sharing secret information and choices with competent agents might jeopardize the very characteristic of secrecy, not doing so would entail the opposite risk of taking ineffective if not harmful decisions.

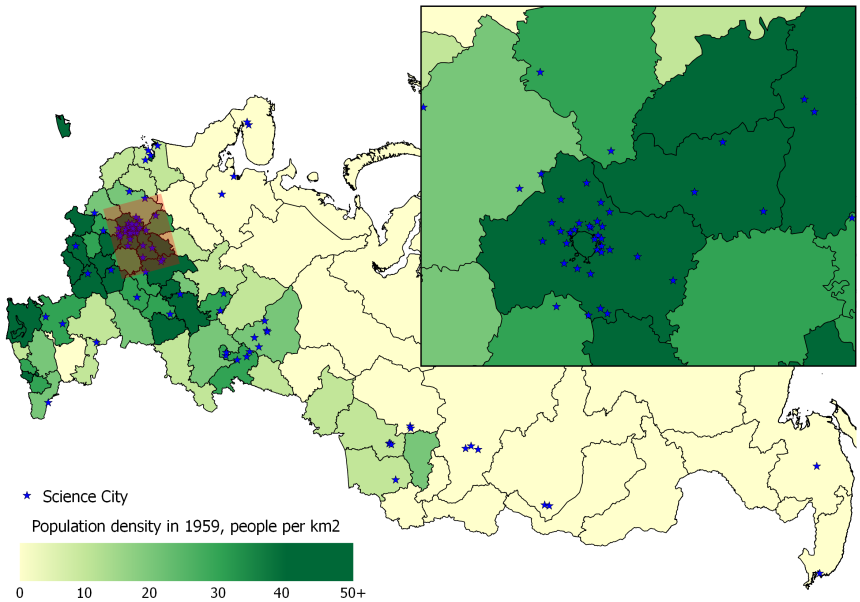
The geographical pattern of Science Cities' locations is described in detail by the Russian geographer Agirrechu (2009). Figure 1 illustrates the location of Science Cities superimposed on a chloropleth map of modern Russian regions, distinguished by their population density. Following Agirrechu (2009, p. 21), Science Cities can be split in two groups of approximately equal size that are identified by their type of location. The first group is composed of those "localities situated in urbanised areas (e.g. in the Moscow region) or within large cities, where the so-called academic towns were organized (e.g. in Novosibirsk, Tomsk, etc.)"[9] – these cities "hosted mainly organizations focusing on theoretical research." By contrast, "[c]ities of the second group were located in the most

up to 1941. In the future, these might allow historians and economists to better reconstruct the process in question.

[8]This is represented by the concept of *planning by feel* by Gregory and Harrison (2005, p. 751). In their writing: "[p]lanners were supposed to distribute materials according to engineering norms, but the first allocations took place before norms were compiled (Gregory and Markevich, 2002; Gregory, 2004). Supply agencies used intuition, trial and error, and *historical experience*. According to one supply official: 'We give 100 units to one branch administration, 90 to another. In the next quarter we'll *do the reverse and see what happens*. You see, we do this on the basis of *feel*; there is *no explanation*' (Gregory and Markevich, 2002, pp. 805–06). According to another: 'Our problem is that we can't really check orders and are not able to check them... We operate partially on the basis of historical material we are supposed to give so and so much in this quarter, and at the same time you are supposed to give us this much.' (cited by Gregory, 2004, p. 172)." Citations and quotes are by Gregory and Harrison, emphasis is ours.

[9]Academic towns were semi-isolated neighborhoods of a larger city, endowed with R&D facilities, housing for R&D staff and their families, as well as basic local infrastructure; the research in natural sciences that was conducted in academic towns was directly linked to the specific issues faced by Siberia. Because they were part of a larger city, we exclude them from our analysis (see also Section 3.2).

Figure 1: Location of Science Cities and regional population density



Sources: Table B.1 and ROSSTAT.

remote areas of the country (although in densely populated regions), far away from large urban centers, highways, industrial facilities, and production fields. The majority of them were surrounded by forests which served as a natural protection from espionage. In these Science Cities, the core enterprises were military-related R&D institutes, design bureaus, pilot plants, and test sites."

This pattern can be easily interpreted as the consequence of a rule of thumb determined by the security-usability trade-off. The specific R&D to be conducted in a perspective Science City would determine the dominant side of the coin; accordingly, Soviet leaders would either choose some remote, highly secretive (but perhaps not too usable) location,[10] or rather an easy to control (but less secluded) place.[11] Agirrechu (2009) also cites other technological and geographical factors that constrained the choice of a location and that, again, depended on the specific R&D specialization: heavy industry and nuclear technology need large amounts of water for their operations, therefore Science

[10]This bears similarities to China's "Third Front" industrialization policy that is examined by Fan and Zou (2017). Military considerations also determined the locations of the "Million Roubles Plants" built in China and studied by Heblich et al. (2018).

[11]The majority of Science Cities of the first group – about one third of the total – are located close to Moscow; according to Agirrechu (2009, p. 21), this is so by virtue of their spatial proximity to "the Academy of Science, the All-Union Academy of Agricultural Sciences, the Academy of Medical Sciences, and some institutes subordinate to ministries."

Cities focused on those areas were typically built close to rivers or lakes; analogously, Science Cities devoted to military shipbuilding and design had to be located on the coast. Furthermore, some Science Cities necessitated timely access to production inputs, and thus had to be placed closer to transportation links, such as railroads.

The locations chosen to host Science Cities were not random, as they typically belonged to selected areas of Russia – more densely populated and urbanized.[12] The quality of our empirical analysis, however, depends on the extent that the chosen locations embed some unobservable factors that made them more (or less) likely to embark on a path of faster demographic and economic development, relatively to other places that were located in the same areas and were otherwise observationally identical to the chosen locations at the time of their selection. It is impossible to provide a definitive response to this question with the currently available archival documentation. Both the historical analysis and the anecdotal evidence, however, clearly point to a negative answer. As for the former, Harrison (2018) cites a series of shortcomings of the decision process taken by Soviet leaders under secrecy: the relevant information was limited to a handful of trusted bureaucrats; these in turn were usually selected on the basis of their loyalty to the régime, rather than on their competence; expert advice was typically absent; decisions were often taken with information limited by other state secrets. In this context, it appears unlikely that choices taken in a planned economy for military and strategic reasons might have been informed by subtle economic factors.[13] The anecdotal evidence, in fact, speaks of very idiosyncratic criteria that often determined the exact locations of certain Science Cities; examples include Sarov and Snezhinsk.[14]

[12]Science Cities are for the most part located in the western, warmer part of Russia, within the humid continental climatic region typified by large seasonal differences in temperature. Historically, the socio-economic development differentials between Russian regions strongly correlate with temperature gradients along a longitudinal axis. In Russia, temperatures in fact change more along the west-east axis than along the north-south axis: for two localities with the same latitude, the eastern one is typically colder.

[13]It is interesting to relate the opinion of Emeritus Professor Mark Harrison (U. of Warwick) on this subject. In personal correspondence with us, he wondered how exactly would the Soviet leaders choose a Science City location, given the aforementioned rule of thumb and constraints. His conclusion is most revealing: "[a]s a child with a blindfold would pin the tail on the donkey, I would guess."

[14]These two places provide a particularly indicative example of idiosyncratic factors affecting the location of Science Cities: sometimes, this was determined by the presence of other Science Cities, or lack thereof. Snezhinsk (Chelyabinsk region) was established as a double of Sarov (Nizhny Novgorod region) with the main purpose of keeping the industry working even if one of the two places were destroyed, but also to create inter-City competition. Since Sarov is located in a relatively remote location in the European part of Russia, Snezhinsk had to be placed in a similarly out-of-reach area, but to the east of Urals. Officials reportedly considered other locations in different regions, but ultimately decided on Snezhinsk because of its proximity to another Science City, Ozyorsk, which could supply inputs to Snezhinsk. This

3 Long-run effects at the municipal level

This section, devoted to the municipal-level analysis, is split into three parts: in the first, we outline our methodology; in the second, we describe the data; in the third, the results.

3.1 Empirical methodology

We compare the long-run outcomes Y_{iq} of municipalities hosting Science Cities with those of other municipalities which were similar in terms of geographical and socio-economic characteristics X_{ik} in the years following the World War II, when the majority of Science Cities were established. $i = 1,\ldots,N$ indexes municipalities; $q = 1,\ldots,Q$ our long-run outcomes of interest; and $k = 1,\ldots,K$ the geographical and historical characteristics we control for. For each long-run outcome, we estimate the Average Treatment Effect on the Treated (ATT), with the treatment being the historical establishment of a Science City in a municipality. To this end, we employ matching techniques.

The central identifying assumption is motivated by our earlier historical discussion. In particular, we assume that the long-run socio-economic outcomes of both Science Cities and their matched counterparts are orthogonal to the treatment, conditional on relevant geographical and historical characteristics and on the paired locations belonging to the same (type of) geographical region. In order to lend further credibility to our approach, we include a number of indicators (which are detailed below) about the military, scientific and economic importance of each municipality in postwar Russia, in addition to matching on cities of equal size and the geographical constraints described by Agirrechu. In our implementation, we take measures to ensure that cities are matched as close in space as possible, especially in the densely populated parts of Russia such as the Moscow region, in the spirit of a "border discontinuity" design. In particular, we include in the conditioning covariates both geographical coordinates (latitude and longitude) and – importantly – the density of historical population and factories within a radius of 50km, thereby controlling for local economic conditions and relieving concerns about the presence of spatially correlated unobservables affecting our results.[15]

pattern of interplay between decisions affecting different Science Cities was not unique; for example, the four places specialized in production of enriched uranium were also located far from each other.

[15] This approach might have some drawbacks: if the "effect" of Science Cities extends beyond their borders, matching two cities too close in space would imply a violation of the Stable Unit Treatment Value Assumption – which likely leads to ATT estimates biased downward. We believe our chosen approach to be more conservative than ignoring spatially correlated unobservables possibly leading to upward biases.

A potential concern with our matching approach is that historical characteristics observed at a particular moment in time are not sufficient to account for the dynamic development path of different localities: a Science City might well look like a matched control municipality at some point in postwar times even if it was already enjoying faster population and economic growth. This would certainly threaten our causal evaluations if pre-trends are likely to continue, and would motivate approaches in the spirit of the "synthetic control matching" (SCM) method, for example by matching on each municipality's time series of urban population. Applying SCM to all our many outcome variables is however unworkable; furthermore, it would not be devoid of shortcomings. As illustrated by Ben-Michael et al. (2019), in fact, SCM is biased with short panel dimensions; they propose an "augmented" SCM through a covariates-based bias correction or equivalently, the inclusion of both covariates and pre-treatment outcomes in a "hybrid" multivariate matching procedure. Inspired by the latter option, we thus implement a Mahalanobis matching algorithm in which the vector of covariates includes the observation of all municipalities' urban population at several points in the past.

More specifically, a Science City s is matched to the ordinary municipality z with the lowest value of the following extended Mahalanobis Distance m_{sz}:

$$m_{sz}\left(\mathrm{x}_{is}, \mathrm{x}_{iz}\right) = \left(\mathrm{x}_{is} - \mathrm{x}_{iz}\right)^{\mathrm{T}} \Sigma^{-\frac{1}{2}} \mathrm{W} \Sigma^{-\frac{1}{2}} \left(\mathrm{x}_{is} - \mathrm{x}_{iz}\right),$$

where x_{ic} is the vector of the K observable covariates for municipality i of type $c \in s, z$; $\Sigma^{-\frac{1}{2}}$ is the Cholesky decomposition of the empirical variance-covariance matrix of the covariates, Σ, while W is a matrix of weights obtained via a "genetic" algorithm aimed at optimizing covariate balance (Diamond and Sekhon, 2013). Matching is performed with replacement, so that a control municipality can be linked to multiple treated cities. We also impose exact matching on selected dummy variables (see section 3.3.1); importantly, we match Science Cities that were subject to the "closed city" status described in section 2, to non-Science Cities that experienced similar restrictions (typically, these are places hosting military bases but lacking an R&D content). In addition, we replicate our analysis using Propensity Score Matching (PSM). However, we find that in our setting PSM is inadequate for guaranteeing that matches are close in geographical space.

We also conducted robustness checks motivated by these concerns by imposing additional constraints on matching – such as a minimal distance between matched observations of 50 or 100km. These exercises, which are available on request, deliver statistically identical results and at the same time slightly worsened covariate balance. In light of this, we prefer not to impose these constraints in our main analysis.

The ATT estimates obtained via PSM (available on request) are usually larger than those obtained in the Mahalanobis case, which we consider more conservative.[16]

In our main analysis, we match Science Cities s to control municipalities z one-to-one, conservatively accepting a higher variance for our estimates in exchange for a lower bias (we also experimented with one-to-many matching, obtaining similar results). We derive a unique association of treated-control observations which is based on the original set of 84 Science Cities in our dataset (see section 3.2.1). However, most ATT estimates are performed on a subset of this matched sample, either because for some Science Cities the information about certain outcomes of interest is not publicly available, or because we remove the current Naukogrady from the analysis. For all our outcomes, we estimate the ATT with and without the correction for the multiple covariates bias, and we perform statistical inference by calculating standard errors based on conventional formulae (Abadie and Imbens, 2006, 2011). Since our coverage of Russian municipalities equals or approximates the universe, we do not apply sampling weights.

3.2 Data and descriptive statistics

We evaluate the long-run effects of Science Cities at the municipal level by employing a unique dataset, which contains information previously unavailable in electronic format. Specifically, it combines: (i) a Science Cities database and (ii) municipal-level data that aggregate various sources of information on historical and current characteristics of Russian cities. Our unit of observation is a Russian municipality;[17] in total, our dataset includes 2,338 such municipalities (the two large cities of Moscow and St. Petersburg are excluded). We used GIS software in order to merge municipal-level and geographical information from different sources. Below, we describe our data and the different sources, introducing the municipal-level variables by type for the sake of clarity. Additional information and references are provided in Appendices B (for the Science Cities database) and C (for the municipal-level information).

[16] Relative to PSM, however, Mahalanobis matching also has some drawbacks: it is known to perform worse with a high number of covariates, or when covariates are not normally distributed (Gu and Rosenbaum, 1993; Zhao, 2004). In order to improve on covariate balance, we calculate Mahalanobis distances taking the logs of those covariates with highly asymmetric empirical distributions. If one such covariate X_{ick} also takes zero values, we use the corresponding quantity $x_{ick} = \log(X_{ick} + 1)$ instead.

[17] In this paper, we use the English term "municipality" to denote the municipal formations (municipal'nye obrazovaniya) of Russia, that is, units at the second administrative level (akin to US counties; also called rayons). We use the word "region" to refer instead to federal subjects (oblast', kray or respublika), that is, units at the first administrative level.

3.2.1 Science Cities Database

The Science Cities database is based on various publicly available sources. Since Science Cities were established in secret, an official and definitive list does not exist; the extant lists are not exhaustive, having been put together following the dissolution of the USSR. Most of the 95 middle-sized urban centers on our list appear in Agirrechu (2009), Lappo and Polyan (2008) and NAS (2002). The database contains information on the location of each Science City, the year the locality was founded, the year in which it became a Science City in the Soviet Union (and the year it became a Naukograd in Russia, where applicable), the type of Science City, whether it was a closed city in the past or is still closed now, and its priority areas of specialization (see Table B.1 in the Appendix). We manually assign Science City status – our treatment – to each municipality; in total, the data include 88 municipalities with at least one Science City.[18] Lastly, we exclude from our analysis four Siberian regional capitals that hosted academic towns (see section 2.2).[19] The reason is that while being very "archetypical" Science Cities, the academic towns in question were incorporated in the municipalities of the respective regional capitals, and we are unable to collect statistical information that is suburb-specific. Hence, keeping these municipalities in the treatment or in the control group would contaminate either. Ultimately, we end up with 84 municipalities hosting at least one Science City.

3.2.2 Municipal-level variables

Socio-economic outcomes. In order to measure differentials in the skill level of local inhabitants, we utilize data from the 2010 Russian census on the overall municipal population, the share of the population whose highest attained education are graduate degrees, and the share of the population that completed any form of postgraduate education.[20] We proxy innovation by the total count of local inventor addresses that appear on patents applied to the European Patent Office (EPO) between 2006 and 2015. Each ad-

[18] NAS (2002) lists four Science Cities for which only their Soviet-era nomenclature is publicly available: Krasnodar-59, Novosibirsk-49, Omsk-5 and Perm-6. Their exact location is still unclear; thus we exclude these four places from the analysis as they cannot be matched to any municipality. In addition, three pairs of Science Cities are located within the same municipalities. Hence, 91 Science Cities are mapped to 88 municipalities with at least one Science City.

[19] These are Irkutsk, Krasnoyarsk, Novosibirsk and Tomsk.

[20] Note that graduate education in Russia refers to achieving a bachelor's or master's degree or their Russian equivalent "specialist," while postgraduate education refers to academic or professional degrees, academic or professional certificates, academic or professional diplomas, or other qualifications for which a graduate education is generally required.

dress is weighted by the inverse of the number of inventors that appear on the relevant patent; we call this measure (local) fractional patents. We divide it by the total number of a city's inhabitants holding a postgraduate qualification to obtain a proxy for average researcher's productivity (average fractional patents).

In addition, we collect information on total employment and per capita wages in construction as well as in the combined R&D-ICT sectors from the Russian Statistical Office (ROSSTAT). Note that ROSSTAT data sources of any kind – like the others detailed below – are typically never available for closed cities, due to national security considerations. Lastly, as accurate GDP data at the municipal level are unavailable in Russia, we use several proxies for economic activity: the night lights intensity (standardized in z-scores) observed by satellites from 1992 through 2011,[21] as well as a number of variables on local SMEs from the 2010 SME census by ROSSTAT. In particular, we examine the overall number, the density and the labor productivity of SMEs, either across all sectors of the economy or specifically in manufacturing.

Budget outcomes. We obtain data on the budgets of Russian municipalities for 2006-16 through ROSSTAT. Once again, the information is missing for all closed cities in the sample. On the revenue side, we are able to differentiate between direct revenues (for example, from local taxes) and transfers from both the federal and regional governments. In addition, we are able to distinguish local expenditures by category, such as education, health care, local infrastructure, and similar. All measures are converted to 2010 prices using ROSSTAT's official CPI indices and averaged over 2006-16.

Amenities. ROSSTAT also allows us – with the mentioned limitations – to access data about certain public goods available in Russian municipalities. Specifically, we calculate the length of local roads that is lit during the night, as well as the number of museums, theaters and libraries in a municipality: services most amenable to the better educated.

Geographical characteristics. We collect or calculate information about several geographical characteristics of Russian municipalities: their area, average altitude, as well as average temperatures in January and July. Since locating close to large amounts of water was necessary for Science Cities of certain specializations, we collect data on each municipality's access to the coast or fresh water (major river or lake).[22]

[21] Night lights can plausibly be used as a proxy for economic activity under the assumption that lighting is a normal good; see Donaldson and Storeygard (2016).

[22] For each municipality, we code this information both as dummy variables (presence or absence of either fresh or salted water within the municipal territory) and as the distance between the municipality's geographical centroid and the closest source of water in question.

Table 1: Municipal-level data, socio-economic outcomes: descriptive statistics

	Science Cities		Other municipalities		*t*-tests on means
	Obs.	Mean (s.d.)	Obs.	Mean (s.d.)	*p*-value
Patent count	84	29.131 (48.830)	2,250	6.025 (160.089)	0.000
Fractional patents	84	11.416 (28.925)	2,250	2.265 (57.396)	0.008
Average patent.count	84	1.837 (4.876)	2,250	0.068 (0.299)	0.001
Average fractional patents	84	0.785 (3.012)	2,250	0.029 (0.107)	0.024
Night lights: 1992-1994	84	1.323 (1.374)	2,250	-0.055 (0.939)	0.000
Night lights: 2009-2011	84	1.505 (1.411)	2,250	-0.062 (0.926)	0.000
Population in 2010 (thousands)	84	95.175 (71.773)	2,250	58.324 (278.488)	0.000
Graduate share in 2010	84	0.224 (0.078)	2,250	0.110 (0.044)	0.000
Postgraduate share in 2010	84	0.006 (0.004)	2,250	0.003 (0.002)	0.000
Employment share in R&D-ICT	67	0.036 (0.034)	2,176	0.007 (0.009)	0.000
Average salary in R&D-ICT (thousands)	67	23.915 (10.299)	2,176	15.366 (7.979)	0.000
Average salary in construction (thousands)	60	23.989 (11.727)	1,479	18.021 (10.473)	0.000
No. of SMEs (thousands) – All	65	1.882 (1.463)	2,140	1.190 (3.116)	0.001
No. of SMEs (per 1000 people) – All	65	24.119 (8.568)	2,159	27.492 (9.617)	0.003
SME labor productivity – All	65	1.615 (0.709)	2,153	0.794 (0.427)	0.000
No. of SMEs (thousands) – Manufact.	65	0.221 (0.186)	2,038	0.120 (0.340)	0.000
No. of SMEs (per 1000 people) – Manufact.	65	0.002 (0.001)	2,038	0.002 (0.001)	0.212
SME labor productivity – Manufact.	63	1.442 (0.713)	2,014	0.768 (0.934)	0.000
Budget: avg. total revenues per capita	64	19.468 (5.984)	2,173	24.761 (51.671)	0.000
Budget: avg. transfers per capita	64	9.538 (3.786)	2,173	18.380 (32.718)	0.000
Budget: avg. tax income per capita	64	9.930 (4.136)	2,173	6.380 (22.790)	0.000
Budget: avg. expenditures per capita	64	19.525 (5.907)	2,173	24.712 (51.011)	0.000

Notes: For brevity, population data split by cohort group, budget expenditures split by category and data about local amenities are not reported; see sections 3.3.4, 3.3.5 and 3.3.6 for mean differences between the treated and control observations. The coverage of data on salaries in construction is incomplete in ROSSTAT. Refer to Appendix C for data sources.

Historical characteristics. We collect information about historical socio-economic characteristics that could affect both Science City status and current outcomes. To account for historical differences in city size and in population density within a controlled radius, we use population data from the first post-World War II census held in the Soviet Union, conducted in January 1959, which provides figures for all urban and large rural localities of that time.[23] To control for pre-trends in population and economic growth as discussed in the previous subsection, we complement this data with the population of Russian cities in 1897, 1926, and 1939 as reconstructed by Mikhailova (2012a,b) using historical census data from each date. Unfortunately, the 1959 census does not provide a population breakdown by educational achievement at the municipal level; to proxy for the pre-existing level of human capital of an urban area we use data on the number of higher education institutions located in a municipality in 1959 (De Witt, 1961), as well as the number of local R&D institutes in 1947 (Dexter and Rodionov, 2016).

To control for the existing level of industrial development in a municipality, we use two pieces of information. The first is the number of the Soviet defense industry plants (factories, research and design establishments) located in each municipality and its surroundings (within 50km) in 1947 (Dexter and Rodionov, 2016). The second is the number of local branches of the State Bank of the USSR in 1946, obtained from its archives. This institution was an instrument of the Soviet economic policy; the geographical dispersion of its branches was indicative of an area's importance for the Soviet development strategies; see also Bircan and De Haas (2020). To account for the fact that some Science Cities needed access to good transportation links, while others had to be located in remote areas far from espionage threats, we use GIS data to measure municipalities' distance from Russian railroads in 1943[24] and from the post-World War II USSR borders.

3.2.3 Summary statistics

Table 1 displays summary statistics for the socio-economic outcomes, while Table 2 provides those for the geographic and historical characteristics, always distinguishing between municipalities hosting Science Cities and all other ordinary municipalities. The tables illustrate that Science Cities were typically located in more populous and warmer places, with a higher historical concentration of industrial plants, universities and R&D institutes. In addition, mean differences between Science Cities and other municipalities are positive and statistically significant for most socio-economic outcomes.

[23]We match those locations to modern municipalities. Note that the population of small rural settlements is not reported in the 1959 census. We account for this by calculating the residual rural population of each Russian region in 1959, and assigning it to the region's municipalities proportionally to their area.

[24]In the Soviet economy, railroads were the workhorse of the transportation network; road transport played only a secondary role (Ambler et al., 1985). Most of the railroads' construction took place in tsarist Russia; even in Soviet times, railroads were important not just for transportation and mobility, but also as drivers of regional industrialization. Following World War II, the Soviet railroad network was further developed under the drive to extract and transport the country's vast natural resources.

Table 2: Municipal-level data, geographical and historical characteristics: descriptive statistics

	Science Cities		Other municipalities		*t*-tests on means
	Obs.	Mean (s.d.)	Obs.	Mean (s.d.)	*p*-value
Is a closed city	84	0.202 (0.404)	2,250	0.011 (0.105)	0.000
Has fresh water	84	0.369 (0.485)	2,250	0.445 (0.497)	0.161
Is mountainous (avg. altitude ≥ 1000km)	84	0.000 (0.000)	2,250	0.040 (0.196)	0.000
Latitude	84	55.698 (3.739)	2,250	53.981 (5.110)	0.000
Longitude	84	47.794 (20.860)	2,250	59.955 (29.410)	0.000
January average temperature (°C)	84	-11.350 (3.699)	2,250	-13.559 (7.045)	0.000
July average temperature (°C)	84	18.528 (1.729)	2,250	18.755 (2.675)	0.251
Distance from the USSR border in 1946	84	669.980 (351.186)	2,250	678.702 (419.358)	0.825
Urban population in 1897 (thousands)	84	3.160 (7.175)	2,250	3.860 (37.469)	0.529
Urban population in 1926 (thousands)	84	5.525 (10.960)	2,250	6.768 (58.439)	0.470
Urban population in 1939 (thousands)	84	16.70 (22.941)	2,250	14.477 (126.641)	0.543
Urban population in 1959 (thousands)	84	34.674 (42.712)	2,250	25.395 (152.375)	0.103
Non-urban population in 1959 (thousands)	84	47.746 (52.838)	2,250	49.590 (153.789)	0.781
Population within 50km in 1959 (thousands)	84	1997.425 (2931.010)	2,250	222.869 (610.306)	0.000
No. of plants in 1947	84	3.310 (4.899)	2,250	2.144 (26.229)	0.131
No. of plants within 50km in 1947	84	369.595 (611.452)	2,250	18.104 (121.580)	0.000
No. of universities in 1959	84	0.143 (0.415)	2,250	0.196 (2.205)	0.409
Has R&D institutes in 1947	84	0.333 (0.474)	2,250	0.055 (0.227)	0.000
(Log) area in km^2	84	5.196 (1.933)	2,250	7.398 (1.741)	0.000
(Log) distance from the coastline	84	6.098 (1.432)	2,250	6.178 (1.246)	0.615
(Log) distance from railroads in 1943	84	1.240 (1.328)	2,250	2.582 (1.953)	0.000
(Log) no. of State Bank branches in 1946	84	0.389 (0.438)	2,250	0.664 (0.388)	0.000

Notes: "Is a closed city," "Has fresh water," "Is mountainous," and "Has R&D institutes in 1947," are coded as dummy variables. Distances are expressed in kilometers. Refer to Appendix C for data sources.

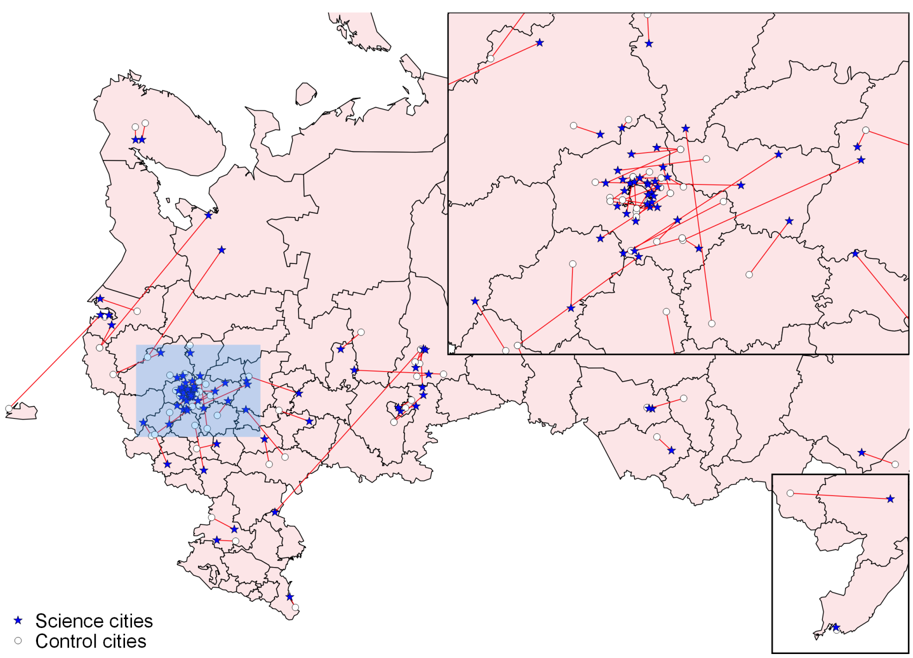
3.3 Empirical results

In what follows, we present the results of the municipal-level empirical analysis, beginning with the description of the matched sample. Following a discussion of our main estimates, we examine a number of additional results for the sake of exploring potential mechanisms that could explain our results.

3.3.1 Quality of matching

The main matching sample consists of 79 municipalities that include a Science City and 61 matched municipalities which do not host any Science City. We impose exact matching on four dummy variables: closed city status, being mountainous, presence of fresh water and of R&D institutes. Out of 84 Science City municipalities in our dataset, 5 are not matched to any control observations,[25] while most control observations are matched to, at most, two Science Cities (a few more in a couple of cases). Figure 2 displays the matched pairs on the map of Russia. Thanks to our choice of covariates, Science Cities and their counterparts are generally matched close in space, especially in the more densely populated and more developed areas of Russia. In particular, municipalities very close to Moscow are typically matched to other municipalities that are also very close to Moscow, which mitigates concerns about the proximity of many Science Cities to the capital of Russia. Table 3 displays the standardized mean difference and the variance ratio between treated and control observations, in both the original and matched sample; it shows that matching achieves a remarkable degree of balance in both the first and the second moment, despite the rigidity of the Mahalanobis algorithm.

Figure 2: Mapping Science Cities and their matches



[25]These are the closed Science Cities of Lesnoy, Seversk, Solnechny, Zelenogorsk and Zheleznogorsk. There are no suitable matches for these cities under our exact matching constraints.

Table 3: Covariate balance: Mahalanobis matching, Science Cities

	Standardized bias		Variance ratio	
	Raw	Matched	Raw	Matched
Latitude	0.383	0.038	0.535	1.143
Longitude	-0.477	0.018	0.503	0.991
January average temperature (°C)	0.393	0.030	0.276	0.775
July average temperature (°C)	-0.101	0.092	0.418	1.106
Distance from the USSR border	-0.023	-0.029	0.701	1.046
Urban population in 1897	-0.026	0.032	0.037	1.030
Urban population in 1926	-0.030	-0.062	0.035	1.040
Urban population in 1939	0.024	0.065	0.033	1.127
Urban population in 1959	0.083	0.032	0.079	0.867
Non-urban population in 1959	-0.016	0.000	0.118	1.054
Population within 50km in 1959	0.838	0.021	23.064	1.060
No. of plants in 1947	0.056	0.062	0.035	0.984
No. of plants within 50km in 1947	0.797	0.014	25.293	1.041
No. of universities in 1959	-0.034	0.059	0.035	0.847
(Log) area in km^2	-1.197	-0.003	1.232	0.800
(Log) distance from the coastline	-0.060	-0.092	1.321	1.289
(Log) distance from railroads in 1943	-0.803	-0.005	0.462	1.147
(Log) no. of State Bank branches in 1946	-0.667	0.000	1.279	1.000

Notes: For each variable listed in the left column, this table reports both the difference in the variance-standardized mean (the "standardized bias," reported in percentage points) and the variance ratio between treated and control observations, for both the raw sample and the matched sample. The matched sample is obtained by "genetic" Mahalanobis matching on the variables above, forcing exact matching on the four dummy variables included in our list of covariates (see the text and the notes to Table 2).

3.3.2 ATT estimation: All Science Cities

The main estimates of the ATT for our 12 outcomes of interest are reported in Table 4. In what follows, we summarize our results, starting from the demographics variables extracted from the 2010 Russian census. Science Cities seem to be, on average, slightly more populated than their matched counterparts, by about 20,000-30,000 people. This difference is driven, for the most part, by the more educated segments of the population. In fact, the share of inhabitants who attained graduate education is about 4-5 percentage points higher in Science Cities; similarly, Science Cities still host more people with some postgraduate qualification (by about 0.2 percentage points). These estimates are substantially smaller than the raw differences, but are generally statistically significant at the 1 per cent level (5 per cent for the total population unadjusted estimate).

Among the innovation outcomes, the absolute fractional patents measure estimate is positive and statistically significant (at the 1 per cent level), like the corresponding measure averaged over postgraduate degree holders (significant at the 5 per cent level).

Table 4: Municipal-level results: Mahalanobis matching, all Science Cities

	Whole sample	*Matched sample* (1 nearest neighbor)				
Outcome	Raw difference	T	C	ATT	ATT b.a.	Γ^* (α = .05)
Population in 2010 (thousands)	35.524*** (9.934)	79	61	24.537** (9.856)	23.992*** (6.616)	2.15
Graduate share in 2010	0.113*** (0.009)	79	61	0.046*** (9.856)	0.045*** (6.616)	4.15
Postgraduate share in 2010	0.003*** (0.000)	79	61	0.002*** (0.000)	0.002*** (0.000)	2.80
Fractional patents	9.060*** (3.398)	79	61	8.938*** (3.125)	8.408*** (3.152)	4.85
Avg. fractional patents	0.755** (0.329)	79	61	0.719** (0.328)	0.710** (0.329)	3.95
Employment share in R&D, ICT	0.028*** (0.004)	67	53	0.025*** (0.004)	0.024*** (0.004)	5.95
Avg. salary in R&D, ICT	8.534*** (1.271)	67	53	5.117*** (1.214)	4.427*** (1.323)	1.85
Avg. salary in construction	5.779*** (1.539)	60	49	8.109** (1.635)	7.171** (1.868)	2.10
Night lights (2009-2011)	1.547** (0.155)	79	61	0.348*** (0.098)	0.220** (0.098)	2.00
No. SMEs, thousands (All)	0.672*** (0.741)	65	51	0.320 (0.331)	0.399** (0.299)	1.65
No. SMEs, thousands (Mnf.)	0.100*** (0.103)	65	51	0.046 (0.065)	0.061** (0.060)	1.75
SME labor productivity (All)	0.813*** (0.088)	65	51	0.206*** (0.071)	0.197*** (0.069)	1.35
SME labor productivity (Mnf.)	0.684*** (0.091)	63	50	0.115 (0.090)	0.119 (0.090)	1.00

Notes: *, ** and *** denote significance at the 10, 5, and 1 per cent level, respectively. Standard errors are reported in parentheses. Raw differences are based on simple dummy variable regressions on the whole sample. In the matched sample, T is the number of matched treated observations; C is the number of matched controls; 'ATT' and 'ATT b.a.' are two estimates of the ATT respectively excluding and including a bias-adjustment term (Abadie and Imbens, 2011). In both cases, standard errors are computed following Abadie and Imbens (2006). Γ^* is the minimum value of parameter $\Gamma \geq 1$, selected from a grid spaced by intervals of 0.05 length, such that in a sensitivity analysis à la Rosenbaum (2002) the set of Wilcoxon signed-rank tests associated with Γ^* do not simultaneously reject the null hypothesis that the outcome variable is not different across the treated and control samples, for tests with α = .05 type I error. Avg. – average; Mnf. – manufacturing.

These results indicate that between 2006 and 2015, Science Cities have applied to the EPO, on average, for 8-10 more fractional patents than their matched municipalities, or about 0.7 more fractional patents for each individual with a postgraduate degree.[26]

[26] We obtain similar results if we use absolute, as opposed to fractional, measures of patent output.

Note that our ATT estimates are statistically indistinguishable from the raw differences for both patent measures, which is arguably because R&D is very spatially concentrated, in Russia and elsewhere. Indeed, ROSSTAT data indicate that high-tech sectors of the economy are more developed in Science Cities, since both measures of employment share and salaries in the combined R&D-ICT sectors register differences that are positive and statistically significant (at the 1 per cent level). In those industries, the share of jobs in these sectors is higher by 2 percentage points in Science Cities, and these jobs pay a monthly salary that is higher by about 5,000 roubles (roughly US$170) in 2010 prices. Interestingly, we observe a similar effect on wages for the construction sector, too.[27]

Lastly, we examine our proxies of overall economic activity. The average of standardized night lights indicators for 2009-11 registers a statistically significant difference in favor of Science Cities, although it appears much smaller than the raw difference and, in the bias-adjusted case, it is only significant at the 5 per cent level. The difference amounts to about 20-35 per cent of the indicator's standard deviation. ROSSTAT's SME census provides a different kind of information. While raw differences suggest that Science Cities are characterized by an overall higher diffusion of SMEs, the corresponding ATT estimates – either relative to all sectors of the economy, or specific to manufacturing – are only significant (at the 5 per cent level) when adjusting for the matching bias. Similar results, which are not displayed in Table 4 for brevity, are obtained for measures of SME density (number of SMEs divided by municipal population). The SME labor productivity ATT estimate is, however, positive and statistically significant at the 1 per cent level, although it is not statistically significant in the case of manufacturing SMEs. In anticipation of our later discussions, we argue that this suggests that the economic effect of Science Cities operates on the intensive (productivity) margin of some industries.

We also perform a sensitivity analysis of our ATT estimates. Following Rosenbaum (2002), we simulate the presence of unobserved factors that would affect both the outcomes and the probability that a municipality hosts a Science City, and we assess to what extent this would influence our conclusions about the presence of statistically significant differences in Y_{iq} between treated and (matched) control observations, for all outcomes $q = 1, \ldots, Q$. The size of the simulated unobserved factor is given by parameter $\Gamma \geq 1$, which measures the hypothesized odds of receiving the treatment ($\Gamma = 1$ is the experimental benchmark). In Table 4 we report, for each outcome variable, the lower

[27] Note that this result is based on a more restricted set of Science Cities, due to the limited coverage of wages in construction in ROSSTAT. Thus, it must be interpreted with caution.

value Γ^* selected from a grid spaced by intervals of 0.05 length that would lead to insignificant Wilcoxon signed-rank tests about differences between treated and control observations.[28] The values of Γ^* are quite high (around 2.5 or more) for the patent outcomes, the night lights measure, the employment share in R&D and ICT and the graduate share. They are satisfactorily high (around 2) for the other demographic outcomes and the two salary measures.[29] These values are in line with our statistical inference about the estimated ATT parameters, and show that our estimates are robust to possible threats to causal identification. Lastly, the values of Γ^* are smaller – between 1.35 and 1.75 – for our SME measures, except for the manufacturing labor productivity measure for which $\Gamma^* = 1$ exactly. Hence, the qualitative conclusions about these outcomes appear less robust, although it must be acknowledged that these inferences suffer from a severely reduced sample size due to ROSSTAT's incomplete coverage.

It must be mentioned that while our results are based on one-to-one matching, the main qualitative conclusions are not altered in the case of one-to-many matching. In fact, increasing the number of matched nearest neighbors usually increases bias in exchange for a reduction in variance, and thus may result in a higher number of (possibly biased) statistically significant ATT parameter estimates. We have obtained similar results by increasing the number of nearest neighbors up to five; we do not present these results here due to space limitations.

3.3.3 ATT estimation: Historical Science Cities

Our interpretation of the estimates in terms of long-run effects would be threatened if, on average, Science Cities still receive a preferential treatment from the Russian government, in the form of direct or indirect support to local R&D or general purpose expenditure, such as infrastructure. In order to assess, to a first degree of approximation, to what extent our results depend on current governmental support, we repeat the above analysis, excluding Science Cities with the official status of Naukogrady in today's Russia. For these 14 Science Cities, the Russian government has resumed the Soviet-era program in

[28]The Wilcoxon signed-rank tests are based on the plain differences between matched pairs for every outcome variable; we select Γ^* as the smallest values of Γ in the grid such that any p-value of the test is higher than $\alpha = .05$. Full results of the sensitivity analysis are available upon request.

[29]$\Gamma = 2$ indicates a simulated unobserved factor that doubles the probability of receiving treatment relative to that of not receiving it, or vice versa; such a high value of Γ would be realistic only in the presence of very serious threats to our conditional independence assumption. Consequently, very high "critical" values of Γ^* associated with a certain outcome – close to 2 or higher – indicate that the results are likely to be robust to such threats.

the early 2000s, although with a less military and more civil focus. We call the remaining Science Cities "historical."

Table 5: Municipal-level results: Mahalanobis matching, "historical" Science Cities

	Whole sample	*Matched sample* (1 nearest neighbor)				
Outcome	Raw difference	T	C	ATT	ATT b.a.	Γ^* ($\alpha = .05$)
Population in 2010 (thousands)	35.524*** (9.934)	65	53	35.712*** (7.635)	33.138*** (5.747)	2.70
Graduate share in 2010	0.113*** (0.009)	65	53	0.038*** (0.006)	0.039*** (0.006)	3.45
Postgraduate share in 2010	0.003*** (0.000)	65	53	0.002*** (0.000)	0.002*** (0.000)	2.35
Fractional patents	9.060*** (3.398)	65	53	3.584*** (0.815)	3.438*** (0.771)	3.50
Avg. fractional patents	0.755** (0.329)	65	53	0.193*** (0.055)	0.187*** (0.057)	2.65
Employment share in R&D, ICT	0.028*** (0.004)	53	45	0.016*** (0.002)	0.014*** (0.002)	3.90
Avg. salary in R&D, ICT	8.534*** (1.271)	53	45	5.282*** (1.144)	3.151** (1.353)	1.70
Avg. salary in construction	5.779*** (1.539)	50	42	8.938*** (1.528)	6.612*** (1.782)	2.45
Night lights (2009-11)	1.547*** (0.155)	65	53	0.372*** (0.084)	0.271*** (0.082)	2.20
No. SMEs, thousands (All)	0.672*** (0.194)	52	44	0.624*** (0.131)	0.666*** (0.110)	1.75
No. SMEs, thousands (Mnf.)	0.100*** (0.024)	52	44	0.086*** (0.017)	0.089*** (0.015)	1.70
SME labor productivity (All)	0.813*** (0.088)	52	44	0.228*** (0.060)	0.136** (0.064)	1.50
SME labor productivity (Mnf.)	0.684*** (0.091)	50	43	0.155** (0.079)	0.054 (0.086)	1.00

Notes: See the notes accompanying Table 4.

The results in Table 5, based on the matched sample restricted to historical Science Cities, are striking. The estimated ATT is, for most outcomes of interest, very similar to the corresponding estimates in Table 4. Statistical inferences and sensitivity analyses à la Rosenbaum generally confirm this assessment.[30] The removal of Naukogrady results in

[30] For a given outcome, the critical Γ^* value is typically smaller in the restricted "historical" subsample. This is due to a reduction in the sample size.

a dramatic change of the estimated effects only for the patent outcomes. The fractional patent count ATT estimate is about 60 per cent smaller than the initial estimates in Table 4, while the average fractional patent measure is about 70 per cent smaller. Nevertheless, both estimates remain significant at the 1 per cent level and robust, as evidenced by a Γ^* from the sensitivity analysis around 3.

The smaller estimated effects on the patent outcomes can be explained in two non-exclusive ways. On the one hand, in an institutional context such as that of Russia, innovation is still predominantly driven by the government sector, and our patent measures reflect the importance of renewed state support to R&D in selected localities. On the other hand, it is possible that in resuming a restricted version of the older Science Cities program, the Russian government has chosen the best former Science Cities for the newer Naukograd program. In either case, we keep observing a positive differential in favor of historical Science Cities for most demographic and economic outcomes of interest. Such differentials are even more surprising as they are clearly independent of the extent to which the government currently supports local R&D, and thus can only be interpreted as long-run effects. Therefore, we find that our initial interpretation of the empirical results is, if anything, reinforced by this restricted analysis.

Finally, we observe that restricting the analysis to the "historical" Science Cities only raises the estimated ATT for the total population by about 40 per cent. Together with the correspondingly lower estimates for the graduate share and the R&D-ICT employment share, this seems to indicate that current Naukograd are inhabited by a lower absolute number of less skilled individuals with respect to the "historical" Science Cities, all else equal. In light of the low number of Naukograd, this apparent finding might be well the result of statistical noise.

3.3.4 ATT estimation: Municipal budgets

We now turn our attention to the analysis of municipal budgets of Science Cities. Its objective is twofold. First, we directly test whether Science Cities, be they historical or current Naukograd, receive a differential amount of direct governmental transfers or local tax earnings (itself a function of local economic activity). In addition, we see this as an opportunity to test the hypothesis by von Ehrlich and Seidel (2018) mentioned in the introduction. They explain their results not by the action of agglomeration forces, but by the persistence of municipal spending in certain, presumably productivity-enhancing, infrastructure. A parallel mechanism could be at work in our setting: for example, since

Science Cities used to be inhabited by relatively more university graduates than other similar localities, their population might have kept a stronger preference for the provision of certain public goods, such as those related to education or even to local physical infrastructure, whose returns are deferred in time.

Russian municipalities collect resources from both local taxes (property taxes, merchant fees, fees for the provision of local services) and from a portion of federal taxes (income tax, business tax and similar) that are paid by local residents. In addition, municipalities receive discretionary transfers from both the federal and the regional governments. In our data, we are able to identify the source of municipal revenues as well as the allocation of expenditures by category (education, health services, local infrastructure and so on) for all Russian municipalities, except closed cities. To obtain relevant measures of interest, we calculate the 2006-16 averages of selected budget items for each municipality and then divide the result by the 2010 municipal population. We estimate the Science City ATT for each of these per capita measures, comparing the fiscal and expenditure patterns of Science Cities with those of their matched counterparts.

Table 6 summarizes these estimates. The table is organized in two panels: the top one (A) reports results relative to all Science Cities for which budget data are available; the bottom one (B) is instead restricted only to the "historical" Science Cities. The two sets of results are similar: in raw differences, Science Cities collect more taxes per capita than ordinary municipalities; however, they receive disproportionately lower total transfers per capita; as a result, both their total revenues and expenditures per capita are smaller. This is only partly mitigated by the fact that Science Cities obtain higher earnings from local taxes. In the matched sample, the ATT estimates for average revenues and expenditures are equal to zero, those for average tax income are positive and statistically significant, and those for average transfers are negative and also statistically significant; all these estimates are pushed towards negative territory upon applying a bias adjustment (especially so within the "historical" subsample). The values of Γ^* are equal or close to 1 for all these outcomes, except for local average tax income in which case Γ^* falls in the range 1.20-1.35 – not the safest result – in both sub-samples.

Our interpretation of these results is based on our understanding of the institutional context: we argue that political economy mechanisms operate to redistribute federal resources in order to achieve approximately similar levels of governmental expenditures per capita across the country. Since Science Cities are typically richer and thus obtain higher local taxes, this often results in lower total transfers in their favor. Governmental

Table 6: Municipal-level results: Mahalanobis matching, municipal budgets analysis

	Whole sample	*Matched sample* (1 nearest neighbor)				
Outcome	Raw difference	*T*	*C*	ATT	ATT b.a.	Γ^* (α = .05)
Panel A: All Science Cities						
Total revenues p.c.	-5.271*** (1.375)	64	51	0.257 (0.887)	-0.298 (1.055)	1.00
All transfers p.c.	-8.738*** (0.869)	64	51	-1.025** (0.478)	-1.159** (0.553)	1.00
Tax income p.c.	3.467*** (0.726)	64	51	1.282** (0.547)	0.862 (0.622)	1.35
Total expenditures p.c.	-5.158*** (1.357)	64	51	0.236 (0.907)	-0.340 (1.070)	1.00
Expenditures on education p.c.	-1.126** (0.512)	64	51	0.849** (0.354)	0.597 (0.446)	1.05
Panel B: Historical Science Cities						
Total revenues p.c.	-5.271*** (1.375)	51	43	-0.584 (0.778)	-3.390*** (1.258)	1.00
All transfers p.c.	-8.738*** (0.869)	51	43	-1.662*** (0.409)	-3.456*** (0.788)	1.00
Tax income p.c.	3.467*** (0.726)	51	43	1.078** (0.488)	0.066 (0.594)	1.20
Total expenditures p.c.	-5.158*** (1.357)	51	43	-0.639 (0.798)	-3.450*** (1.256)	1.00
Expenditure on education p.c.	-1.126** (0.512)	51	43	0.479* (0.286)	-0.679 (0.512)	1.00

Notes: See the notes accompanying Table 4; p.c. – per capita.

support for Science Cities may also exist in the form of direct expenditures appearing only in the federal budget: unfortunately, such data are not available to us. Yet, if Science Cities were still strategically important for the federal government, we would expect – if anything – to observe less of a symmetry between revenues and transfers per capita. In other words, the government may want to complement direct intervention with more indirect subsidies. We do not observe this in the data.

To test the vES hypothesis, we examine whether Science Cities still differ from their matched localities in terms of per capita expenditures on a number of entries of their municipal budget. For brevity, we only report results on the "education" entry, as we do not observe any significant or otherwise interesting differences for other entries. Our initial hypothesis is that in Science Cities, a more educated local population might have demanded stronger investment in education for their children – which would have explained persistent local advantages. However, the data lend little empirical support to this idea. While Science Cities indeed seem to spend on average more on education, this difference is (weakly) statistically significant only before applying the bias adjustment in the larger sample; in addition, the rank tests from the sensitivity analysis suggest that this result is not robust. Note that in Russia, the educational system is predominantly public and highly centralized; municipal expenditures in education are mostly related to the maintenance of the local schools.

3.3.5 ATT estimation: Amenities

The previous analysis of municipal budgets regarded differences in the flows of spending in public goods. But what if Science Cities and their matched counterparts differed in the stocks instead? It is plausible that Science Cities, which were shaped for the purpose of accommodating researchers and scientists, may have inherited a larger endowment of cultural public goods, such as libraries, museums and theaters – usually most appreciated by the better educated (especially so in a cultural setting such as the Russian) – or other types of public goods. This in turn might explain the higher density of skilled workers in former Science Cities. We explore this hypothesis using our ROSSTAT data; our estimates are reported in Table 7.

The results are virtually identical across the two panels. To our surprise, former Science Cities appear to possess, if anything, fewer cultural amenities than ordinary municipalities. This difference, however, becomes not statistically significant in the matched sample, except for libraries and for museums in the "historical" subsample. Once again, we believe there are redistributive factors at play, perhaps inherited by the emphasis that the communist ideology placed on shared education and culture. Regardless, this specific channel is unlikely to explain the persistent human capital advantages of former Science Cities. Where these places seem to do better than their matches is in terms of nocturnal illumination of their roads, measured as share, over the total length of roads, that is lit in the night. This finding is not too surprising, since the state and maintenance of roads and other terrestrial transportation links is a heartfelt issue in Russia because of geographical and institutional reasons. Nevertheless, the low Γ^* attached to this measure suggest that the strongly significant ATT estimates are perhaps the consequence of a few outlier observations; thus, this is hardly a major driver of our main results.

Table 7: Municipal-level results: Mahalanobis matching, amenities

	Whole sample	*Matched sample* (1 nearest neighbor)				
Outcome	Raw difference	T	C	ATT	ATT b.a.	Γ^* ($\alpha = .05$)
Panel A: All Science Cities						
Libraries per 1,000 inhabitants	-0.601*** (0.029)	64	51	-0.091*** (0.030)	-0.066** (0.032)	1.00
Museums per 1,000 inhabitants	-0.029*** (0.003)	64	51	-0.004 (0.004)	-0.003 (0.004)	1.00
Theaters per 1,000 inhabitants	0.002*** (0.001)	64	51	-0.004 (0.004)	-0.003 (0.004)	1.00
Share of streets lit in the night	0.267*** (0.034)	64	51	0.107*** (0.034)	0.107*** (0.037)	1.20
Panel B: Historical Science Cities						
Libraries per 1,000 inhabitants	-0.601*** (0.029)	51	44	-0.114*** (0.031)	-0.115*** (0.037)	1.00
Museums per 1,000 inhabitants	-0.029*** (0.003)	51	44	-0.007* (0.004)	-0.008** (0.004)	1.00
Theaters per 1,000 inhabitants	0.002*** (0.001)	51	44	0.000 (0.001)	0.000 (0.001)	1.00
Share of streets lit in the night	0.267*** (0.034)	50	44	0.134*** (0.033)	0.104*** (0.040)	1.30

Notes: See the notes accompanying Table 4.

3.3.6 ATT estimation: Demographic dynamics

Another concern is that our results may not be long-lasting. In our conceptual framework (Section 4) we postulate the existence of "persistence forces," independent of other endogenous mechanisms, that induce path-dependence from the Soviet-era allocation of the labor force; say, for example, that the latter affects individual preferences for locations. In the real world, however, workers are slowly replaced by younger workers from newer generations. If new generations do not share the preferences or the characteristics of their ancestors, spatial equilibrium can over time lead to mean reversion – even in the presence of agglomeration forces, thanks to the action of random shocks. This feature is typical of empirical studies in economic geography, perhaps most famously that by Davis and Weinstein (2002) on post-World War II Japan. In such a scenario, our results could not be interpreted as true long-run effects, but rather as snapshots of a long transition back to a steady state.

We investigate the possibility that the advantage of Science Cities wanes over time by exploiting additional information available in our dataset. Specifically, the Russian census data allow us to identify the number of residents in each municipality by type of attained education within each cohort of birth. This lets us assess to what extent our results on urban educational levels are driven mainly by older cohorts, or instead substantially depend on younger cohorts as well. To this end, we split the population of each municipality, as it is observed in the 2010 census, between the "young" (those born after 1965), and the "old" (those born on or before 1965). At the time of the dissolution of the USSR (1991-1992), the older individuals in the "young" group who had obtained a university degree were starting their professional careers and presumably could move more easily. Furthermore, those who were underage at the time of the transition might have pursued less education than their ancestors (mean reversion). Both factors would predict a more equal distribution of young graduates between Science Cities and their matched counterparts.

Table 8: Municipal-level results: Mahalanobis matching, "dynamic" analysis

	Whole sample	*Matched sample* (1 nearest neighbor)				
Outcome	Raw difference	T	C	ATT	ATT b.a.	Γ^* ($\alpha = .05$)
Panel A: All Science Cities						
Graduate share: born ≤ 1965	0.122*** (0.010)	79	61	0.055*** (0.008)	0.054*** (0.009)	4.25
Graduate share: born > 1965	0.107*** (0.008)	79	61	0.034*** (0.007)	0.032*** (0.007)	2.30
Postgraduate share: born ≤ 1955	0.003*** (0.001)	79	61	0.002*** (0.001)	0.002*** (0.001)	1.85
Postgraduate share: born > 1955	0.003*** (0.000)	79	61	0.002*** (0.000)	0.002*** (0.000)	2.35
Panel B: Historical Science Cities						
Graduate share: born ≤ 1965	0.112*** (0.010)	65	53	0.042*** (0.007)	0.046*** (0.007)	3.30
Graduate share: born > 1965	0.107*** (0.008)	65	53	0.030*** (0.008)	0.030*** (0.009)	2.10
Postgraduate share: born ≤ 1955	0.003*** (0.001)	65	53	0.002*** (0.000)	0.002*** (0.000)	1.60
Postgraduate share: born > 1955	0.003*** (0.000)	65	53	0.002*** (0.000)	0.002** (0.000)	2.00

Notes: See the notes accompanying Table 4.

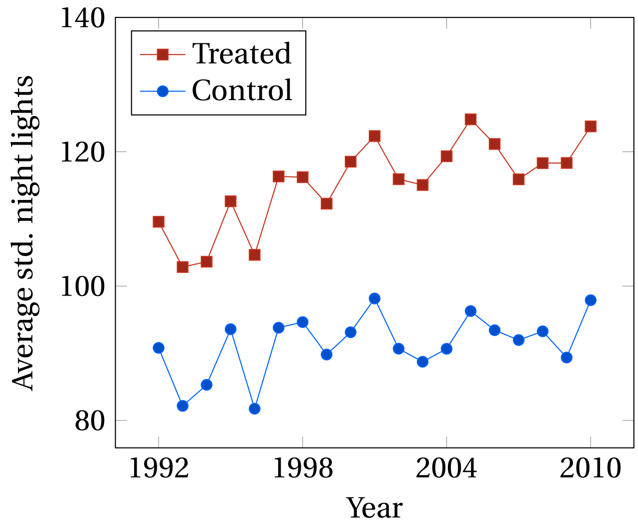
Using our matched sample, we estimate the Science Cities ATT for the graduate share of the population separately for the "old" and "young" groups. The results in Table 8 show that while the differences are indeed larger for the older group, they are positive and statistically significant for the younger one as well, albeit amounting to about 60 per cent of the older group's. If current Naukogrady are removed from the sample there is even less of a difference. In the case of the postgraduate share, we define the threshold year of birth as 1955, taking into account the fact that in Russia, postgraduate education is characterised by a long average duration; the results are qualitatively similar[31] and are not sensitive to the choice of the threshold. Thus, this analysis provides little evidence in favor of the mean reversion hypothesis: it appears that the children of Soviet inhabitants of Science Cities pursue educational and locational choices that are largely similar, albeit not identical, to those of their ancestors.

3.3.7 ATT estimation: Economic dynamics

A logical next step would be to assess mean reversion in economic outcomes. If the relative skill level of Science Cities and that of comparable municipalities are equalised over time, we would expect economic convergence as well. Our data do not allow us to track the evolution of our proxies of economic activity over the post-transition years, with the exception of the night lights measures. In Figure 3, we plot the average of the standardized night lights indicator separately for Science Cities and their matched controls, for every year from 1992 to 2010. Note that while the two groups share parallel annual fluctuations, Science Cities appear to constantly outperform their counterparts, with hardly any catch-up by the control group. However, this observation may also be due – albeit unlikely – to an extreme path dependence of some random unobserved factors that are not explained by Science City status. To clear this concern, we perform some formal regression-based tests, allowing for temporal persistence in the unobservable factors driving each municipality's night light measures. These tests are relegated to Appendix D; in summary, the effect of Science Cities is statistically robust and constant over time, offering once again little support to the mean-reversion hypothesis.

[31] We observe a secular increase in the attainment of postgraduate education in Russia following the transition, which is opposite to the general trend observed for tertiary education. Among all municipalities, the unweighted average share of graduates in the old group is about 12.5 per cent, while it amounts to about 11.0 among the younger (24.5 per cent versus 21.5 per cent in Science Cities). Conversely, the postgraduate share is 0.15 per cent in the old group and 0.33 per cent in the young group (0.50 per cent versus 0.63 per cent in Science Cities).

Figure 3: Time series of the average standardized night lights indicators, 1992-2010



3.3.8 Robustness: effects by categories of Science City

Finally, we investigate whether our results are mostly driven by specific types of Science Cities, or by observable characteristics of Science Cities other than their possible current Naukograd status. In particular, we take into account the following categorizations. Reminiscent of the security-usability trade-off discussed in section 2.2, we first ask ourselves whether Science Cities originally serving top secret military purposes – that is, those in aviation/rocket or in nuclear science – have embarked on a different pattern of economic development relative to other Science Cities whose R&D had a less strategic content. One can hypothesize that, for example, the former group had benefited from larger investment in the past, whose effect would be visible to the present day. Second, we look at Science Cities that were built from scratch vis-à-vis those created out of pre-existing settlements; one may conjecture that the pattern of demographic and economic development has been different in the two groups.[32] Third, we examine whether the long-run outcomes of the Science Cities identified by Agirrechu (2009), whose list we consider incomplete and best amended through our other cited sources, are in any way different from those of the residual cities in our list.[33]

[32] For example, Science Cities built from scratch may have initially grown faster, while stagnating following the transition to a market economy in case they had been placed in inconvenient locations. Note that Science Cities built from scratch were settlements of new urbanization located in a rayon (municipality) which included other, possibly sparsely populated, settlements; under our matching approach, these Science Cities are paired to municipalities with similar characteristics and pre-trends.

[33] The original list by Agirrechu (2009) includes 75 Science Cities, while our extended list counts 95.

To provide a first degree assessment about differences between the cited categories, we perform t-tests about the means of the treated-control differences in the matched sample. Given the small size of some of the subsamples in question, we restrict our attention to some key outcomes for which our information is complete for all cities. Specifically, we look at the total number of patents, the night lights measure averaged over 2009-11, total population in 2010, and the share of that population with at least a graduate degree. We summarize the results of our t-tests, for each categorization and key outcome, in Table 9. In addition, Appendix E provides a visual representation of the treated-control differences, split by group, for all of these cases.

Table 9: Robustness t-tests: differences between types of Science Cities

Criterion	Outcome	p-value
"Secrecy" vs. "usability"	Total patents	0.962
	Night lights (2009-11)	0.136
	Total population	0.122
	Graduate population share	0.063
Built from scratch vs. others	Total patents	0.421
	Night lights (2009-11)	0.623
	Total population	0.297
	Graduate population share	0.736
Agirrechu's list vs. others	Total patents	0.004
	Night lights (2009-11)	0.111
	Total population	0.747
	Graduate population share	0.585

Notes: The p-values on the third column are obtained from t-tests on the mean treated-control differences between two groups identified by the criteria listed in the first column. The differences are measured for every outcome listed in the second column. Appendix E provides a visual representation of the quantities being tested.

By inspecting the twelve p-values reported in Table 9, it appears that only one t-test decidedly rejects the null hypothesis. Specifically, the mean difference in total patent output between Science Cities in Agirrechu's list and the others is statistically significant at the 1 per cent level, which is unsurprising in light of the results from section 3.3.3 and the fact that all current Naukogrady are in Agirrechu's list. The test about the graduate share for the "secrecy" vs. "usability" split rejects the null hypothesis of no difference at the 10 per cent level, a piece of evidence that we deem inconclusive (it may well be a Type I error). All other null hypotheses cannot be rejected under conventional confidence levels; most of the p-values are notably large – including the demographic outcomes for the "built from scratch" split. Thus, we conclude that the specific qualities of Science Cities we have examined are hardly major drivers of our main results.

4 Model and Mechanisms

To facilitate the interpretation of the long-run effect of the Science Cities program, we estimate a spatial equilibrium model which is typical in urban economics. Specifically, we adapt the model by Moretti (2011, 2013) which itself extends Rosen (1979), Roback (1982) and Glaeser and Gottlieb (2008, 2009). We first describe the model and its equilibrium properties; then, we discuss the resulting estimates and their interpretation.

4.1 Model set-up

Consider a set $\mathscr{C}$ of Russian cities, of dimension C and indexed as $c = 1, \ldots, C$. These cities could be inhabited by two different types of workers: those of high educational level or "skill," and those of relatively lower skill. This dichotomous classification is conventionally interpreted in terms of differences in higher educational achievement. In this context highly skilled workers can be alternatively identified, more narrowly, as researchers engaged in R&D – typically a subset of all university-educated individuals – while low-skilled workers represent all other workers in the remaining sectors. The model is general enough to allow for both interpretations. We denote the mass of highly skilled workers employed in city c at time t as H_{ct}, while L_{ct} is the corresponding notation for low-skilled workers. We also use lower-case letters, respectively h_{ct} and ℓ_{ct}, to denote the logarithm of such masses.

At time $t = 0$, all cities are part of the Soviet Union which, for idiosyncratic reasons, attributes the status of Science City only to a subset $\mathscr{S} \subset \mathscr{C}$ of them (write its complement as $\mathscr{Z}$, with $\mathscr{S} \cup \mathscr{Z} = \mathscr{C}$). We interpret the Science Cities program primarily as a spatial reallocation of workers according to skill status in the context of a planned economy. Thus, the Soviet Union allocates proportionately more highly skilled workers to a Science City $s \in \mathscr{S}$ than to an ordinary locality $z \in \mathscr{Z}$, so that $(H_{s0} - H_{z0}) > 0$. We make no statement about the sign of the corresponding difference for the low skilled, $(L_{s0} - L_{z0})$: it may be positive, as R&D workers needed local supporting personnel and services, or negative, which would reflect the spatial segregation of economic activity in the USSR.

At time $t = 1$, all cities are part of modern Russia, a market economy, and workers of both types self-select into either location. Combining Moretti (2011, 2013) with the typical characterization of random utility models, we express the logarithmic indirect utility u_{nic} of an individual i of type $n \in h, \ell$ living in city c, as a linear function of wages, local characteristics such as amenities, and idiosyncratic preferences:[34]

$$u_{nic} = w_{nc} + a_{nc} + e_{nic}, \tag{1}$$

where w_{nc} is the log-wage earned by workers of type n in city c; a_{nc} represents factors specific to city c (such as amenities or public goods) that make it a more enjoyable location to live in, and whose specific value might vary by skill group; while e_{nic} denotes the idiosyncratic taste of individual i for city c, which is assumed to be a random shock following a Type I Extreme Value (Gumbel) distribution, with zero location parameter and unitary scale parameter as per the standard normalization of random utility models.

Lastly, to close the model we introduce two types of firm: that which employs highly skilled workers, and that which relies on lower skilled workers. While in Moretti's analysis this was largely a simplification meant to abstract from the degree of substitutability between skills, this characteristic of the model can be given a contextual interpretation here: if workers of type h are researchers, type h firms correspond to the R&D sector, while type ℓ firms represent the rest of the local economy. The log-output y_{nc} of type n firms in city c is determined according to a Cobb-Douglas technology:

$$\begin{aligned} y_{hc} &= g_{hc} + \theta_h h_c + \mu h_c + (1-\mu)\, k_{hc} \\ y_{\ell c} &= g_{\ell c} + \theta_\ell h_c + \mu \ell_c + (1-\mu)\, k_{\ell c}, \end{aligned} \tag{2}$$

where g_{nc} is the city- and type-specific total factor productivity, while k_{nc} is the log-capital employed by the firms of type n in city c. The supply of capital is infinitely elastic and its cost is the same for all firms in the two cities s and z. For simplicity, the elasticity of labor is equal to $\mu \in (0,1)$ for both types of firm in both cities. Note that firms of type ℓ do not hire workers of type h, but take h_c as given.

[34]Typically, in these models workers' utility also depends – negatively – on city-specific price indexes r_c. For simplicity, we assume that local prices are identical in the two locations: $r_z = r_s$. If r_c represents rents, this could follow if houses are supplied completely elastically in two competitive markets employing the same technology. We also abstract from congestion effects à la Glaeser and Gottlieb (2008, 2009) and any kind of negative externalities that may depend on a city's population. These simplifications allow us to focus our discussion on the interplay between labor supply and agglomeration effects.

The interpretation of parameters $\theta_h \geq 0$ and $\theta_\ell \geq 0$ is as follows. For type h firms, $\theta_h > 0$ allows for increasing returns due to knowledge spillovers; since the productivity of highly skilled workers grows more than proportionately to their number, this introduces an agglomeration force in the economy. Note that $\theta_h = 0$ implies constant returns to scale in type h firms. If knowledge spillovers also operate between firms, and the size of the local highly-skilled workforce can affect the productivity of the less skilled workers as well, then $\theta_\ell > 0$. The model provides different equilibrium predictions to the extent that $\theta_h > 0$, $\theta_\ell > 0$, or both. Throughout the analysis, we assume that high-to-high spillovers are small enough relative to the baseline elasticity of labor inputs: $\theta_h < \mu$. This rules out degenerate or otherwise counter-intuitive solutions.

4.2 Spatial equilibrium

In a spatial equilibrium at $t = 1$, workers choose freely in which location to reside. As is typical in models of this literature, the equilibrium is determined by the interplay of agglomeration and "divergence" forces that lead to spatial dispersion. In this model, the only divergence force is given by individual preferences. In what follows, the spatial equilibrium is characterized by comparing the outcomes of any two cities in a pair. To better introduce the next section about estimation, here we illustrate the equilibrium equations in terms of the outcomes of a Science City $s \in \mathscr{S}$ and an ordinary municipality $z \in \mathscr{Z}$; however, the results that follow extend to any pair of cities in $\mathscr{C}$. Furthermore, we drop for convenience all time subscripts referring to $t = 1$.

In equilibrium, the mass of workers of a certain type that choose to live in a certain city is proportional to the probability that one individual of that type chooses that location. By the standard properties of logit-type multinomial choice models, one obtains:

$$h_s - h_z = \log \frac{H_s}{H_z} = (w_{hs} - w_{hz}) + (a_{hs} - a_{hz}). \tag{3}$$

an equation interpreted as the labor supply of highly-skilled workers between the two cities s and z. In the labor market, the equilibrium wage differentials $(w_{hs} - w_{hz})$ are obtained as the difference between the marginal productivity of highly skilled labor in the two cities; this difference, in turn, depends on the equilibrium in the capital market.[35] A

[35]Equilibrium in the capital market implies that the marginal productivity of capital must be equal in the two cities: $\mu(k_{hs} - k_{hz}) = (\theta_h + \mu)(h_s - h_z) + (g_{hs} - g_{hz})$. As a consequence, the difference between the inverse labor demands in the two cities can be expressed as: $(w_{hs} - w_{hz}) = \mu^{-1}\left[(g_{hs} - g_{hz}) + \theta_h (h_s - h_z)\right]$.

symmetric analysis applies to the case of low-skilled labor. As a result, the relative difference in equilibrium highly skilled employment between the two cities can be expressed as follows:

$$(h_s - h_z) = \left(\frac{\mu}{\mu - \theta_h}\right)\left[\frac{1}{\mu}\left(g_{hs} - g_{hz}\right) + (a_{hs} - a_{hz})\right]. \tag{4}$$

Equation (4) is easily interpreted. There are two forces that cause city s to host a larger number of researchers and highly skilled workers after the transition relative to city z. These are: (i) inherent productivity differentials ($g_{hs} - g_{hz} > 0$) and (ii) superior amenities ($a_{hs} - a_{hz} > 0$) both in favor of city s. All these forces are stronger the larger are the agglomeration effects (higher θ_h). Importantly, agglomeration effects alone are not sufficient to cause employment differentials, at least in the equilibrium under analysis: they only complement the two factors that affect the supply of high skilled labor. Agglomeration externalities also positively affect the equilibrium wages of highly skilled workers. By combining the above equilibrium equations, one obtains:

$$\begin{aligned}(w_{ns} - w_{nz}) &= \left(y_{hs} - y_{hz}\right) - (h_s - h_z) \\ &= \frac{1}{\mu - \theta_h}\left(g_{hs} - g_{hz}\right) + \frac{\theta_h}{\mu - \theta_h}(a_{hs} - a_{hz}).\end{aligned} \tag{5}$$

With equal levels of TFP ($g_{hs} = g_{hz}$) wage and productivity differentials can still be sustained by agglomeration forces ($\theta_h > 0$) if amenity differential persist: $(a_{hs} - a_{hz}) > 0$.[36]

We complete the description of the model by stating the results for the low-skilled workers, taking as given the equilibrium for the highly skilled as expressed in (4). For the low-skilled, the equilibrium differentials in wage and productivity are obtained as:

$$(w_{\ell s} - w_{\ell z}) = \left(y_{\ell s} - y_{\ell z}\right) - (\ell_s - \ell_z) = \frac{1}{\mu}\left(g_{\ell s} - g_{\ell z}\right) + \frac{\theta_\ell}{\mu}(h_s - h_z), \tag{6}$$

which again may be sustained by high-to-low spillovers ($\theta_\ell > 0$) even with equalized TFP. Given (6), the equilibrium difference in low-skilled employment reads as follows.

$$(\ell_s - \ell_z) = \frac{1}{\mu}\left(g_{\ell s} - g_{\ell z}\right) + \frac{\theta_\ell}{\mu}(h_s - h_z) + (a_{\ell s} - a_{\ell z}). \tag{7}$$

If $\theta_\ell > 0$, the low-skilled population is a positive function of the high-skilled one.

[36]Note that these results would still apply, in qualitative terms, if rents or congestion effects were allowed to vary by city and to depend on a city's total population. In this case real wage differentials would be smaller than nominal wage differentials, thus restraining labor mobility in equilibrium.

4.3 Estimation of the model

The equilibrium predictions of this model are well suited to empirical estimation. In what follows, we develop an econometric analysis of the spatial equilibrium predictions that is restricted to the matched sample described in section 3.3.1. The objective of this exercise is to quantify some key parameters of the model to aid the interpretation of the main results from the municipal-level matching analysis. The unit of observation from this analysis is a matched pair of cities: in the remainder of this section, the indices s and z denote respectively a Science City and its matched counterpart from our main matched sample analyzed in section 3.

We begin by developing the expressions of municipal-level TFP and amenities in terms of observable characteristics, as follows. For $n = h, \ell$, we write:

$$a_{ns} - a_{nz} = \rho_n + \sum_{k=1}^{K} \kappa'_{kn} (x_{ks} - x_{kz}) + \sum_{p=1}^{P} \delta'_{pn} \left(b_{ps} - b_{pz}\right) + \omega^a_{nsz} \tag{8}$$

$$g_{ns} - g_{nz} = \varphi_n + \sum_{k=1}^{K} \xi'_{kn} (x_{ks} - x_{kz}) + \sum_{p=1}^{P} \zeta'_{pn} \left(b_{ps} - b_{pz}\right) + \omega^g_{nsz}, \tag{9}$$

where x_{kn} is one of the K observable geographical and historical characteristics that we match upon, for $k = 1, \ldots, K$, while b_{pn} is one of P budget items (with $p = 1, \ldots, P$) that are observed in our municipal budget data (such as expenditures in education, physical infrastructure, utilities, security, etc.). In addition, ω^a_{nsz} and ω^g_{nsz} are two error terms. Under this specification, the primitive drivers of the spatial equilibrium are made dependent on observable factors. It is natural to conjecture that the desirability of a certain location, or the productivity of its firms, is a function of geographical characteristics as well as of local public goods. To facilitate identification, we assume that individuals do not internalize the TFP shocks ω^g_{nsz} at the time of taking their location decisions, effectively treating these as unexpected productivity shocks.

While the interpretation of the parameters expressed as κ'_{kn}, δ'_{pn}, ξ'_{kn} and ζ'_{pn} is obvious, the two constants from (8) and (9) deserve more elaboration. Parameter ρ_n measures the average extra desirability, for workers of type n, of a Science City relative to an ex-ante otherwise identical location. This parameter can be also dually interpreted as an implicit cost: for example, the cost associated with intra-city mobility (in Russia, internal mobility used to be very costly due to regulation inherited from Soviet times[37]).

[37]A system of internal visas was in place until the early 2000s: (the so-called propiska system). Studies

In light of the log-utility expression (1), the numerical value of ρ_n can be evaluated in terms of units of log-wage, indicating the minimal compensating variation that the average worker of a given type would accept in order to leave a Science City. We interpret this parameter as a grand measure of the persistence forces operating in our model.[38] Parameter φ_n is instead more easily understood as the extra productivity benefit of locating in a Science City for firms of type n. This advantage might depend on factors unaccounted by the model, such as the level of social trust or better local institutions.

By substituting (8) and (9) into selected relationships that define our spatial equilibrium model – namely (4), (7) and the two labor market equilibrium equations – we obtain the following system of simultaneous equations:

$$h_s - h_z = \frac{\mu\rho_h + \varphi_h}{\mu - \theta_h} + \sum_{p=1}^{P} \delta_{ph}\left(b_{ps} - b_{pz}\right) + \omega^d_{hsz} \tag{10}$$

$$\ell_s - \ell_z = \rho_\ell + \frac{\varphi_\ell}{\mu} + \sum_{p=1}^{P} \delta_{p\ell}\left(b_{ps} - b_{pz}\right) + \frac{\theta_\ell}{\mu}\left(h_s - h_z\right) + \omega^d_{\ell sz} \tag{11}$$

$$w_{hs} - w_{hz} = \frac{\varphi_h}{\mu - \theta_h} + \sum_{p=1}^{P} \zeta_{ph}\left(b_{ps} - b_{pz}\right) + \frac{\theta_h}{\mu}\left(h_s - h_z\right) + \omega^w_{hsz} \tag{12}$$

$$w_{\ell s} - w_{\ell z} = \frac{\varphi_\ell}{\mu} + \sum_{p=1}^{P} \zeta_{p\ell}\left(b_{ps} - b_{pz}\right) + \frac{\theta_\ell}{\mu}\left(h_s - h_z\right) + \omega^w_{\ell sz}, \tag{13}$$

where, for $n = h, \ell$, δ_{pn} and ζ_{pn} are functions of δ_{pn},$'$ and ζ'_{pn} and other parameters; while ω^d_{nsz} and ω^w_{nsz} are functions of ω^a_{nsz}, ω^g_{nsz} as well as of the differences in the geo-historical characteristics $x_{ks} - x_{kz}$. In fact, the latter are not included in the system: since the analysis is restricted to the matched sample, they are expected to have mean zero and best treated as disturbances; given the small sample size, including them in the

about internal migration rates in Russia in the 1990s show that they were very low (Andrienko and Guriev, 2004; Friebel and Guriev, 2005).

[38]One could microfound the differential expressed as ρ_n by conjecturing that it is a function of the initial allocation of workers at $t = 0$. Write, for example:

$$\rho_h = r\left(H_{s0} - H_{z0}\right)$$
$$\rho_l = r\left(L_{s0} - L_{z0}\right),$$

where $r(\cdot)$ is an increasing monotone function with $r(0) = 0$. This assumption introduces a mechanism of path-persistence: if some individuals used to reside in a specific city during Soviet times, they are likely to prefer to stay there. Consequently, the average bias of workers of a given type depends on their relative allocation at $t = 0$. In future research, we plan to collect data about the population of pre-transition Russia – split by the level of education – in order to explore this possibility in more detail.

estimation would overfit the model. Another implication of restricting the analysis to the matched sample is that the identification of the constant parameters ρ_n and φ_n (for $n = h, \ell$) is based on a difference-in-differences approach, where identification and the parallel trend assumption rest on the matching strategy from Section 3. The model can be expanded further; an extension that we examine in our estimates is a split of the demographic outcomes equations (11)-(10) by age group: young and old, as described in section 3.3.6. In this case, we estimate one ρ_n parameter for each age-skill group.

Before moving to the empirical results obtained from the estimation of this model, a few brief observations about the econometrics are in order. First, the baseline labor share μ is not identified as we do not observe data about capital at the municipal level; for this reason, we calibrate it as $\mu = 0.6$; this value is supported by the typical production function estimates found by Kuboniwa (2011) in an econometric study about Russian firms. Instead, the parameters ρ_n and φ_n (for $n = h, \ell$) are identified residually, once all other right-hand side variables are netted out; this is consistent with their interpretation. In addition, the right-hand-side variable $(h_s - h_z)$ is unlikely to be mean-independent of the error term in (11), hence it is not treated as an instrument for that equation in the moment conditions; yet parameter θ_ℓ is identified in the low-skilled wage equation (13).

Next, observe that the error terms of the model are by construction correlated across the equations of the system, as well as between two dyads that share the same matched control municipalities z. Therefore, we jointly estimate the model via GMM, thereby allowing for arbitrary correlation between the error terms of the various equations; in addition, we cluster standard errors by small groups (usually comprising two dyads) that share the same control location. In order to improve on the statistical performance of the model, we do not impose the restrictions implied by the model on the parameters δ_{pn} and ζ_{pn} for the municipal budget entries. Regarding the sample size, most equations are estimated using all 64 pairs of observations in the matched sample such that the information about both salaries in R&D-ICT sector and the budget entries b_{pc} are available in ROSSTAT for both municipalities s and z in the dyad. The low-skilled wage equation (13), however, is estimated with a subset of those observations – 45 in total – because of ROSSTAT's incomplete coverage of salaries in the construction sector.

With these caveats in mind, we move to the description of our estimates, which are reported in Table 10. In column 1 we report the results from the estimation of a bare-bones version of our model, where the extra TFP effects φ_n and the parameters δ_{pn} and ζ_{pn} associated with the municipal budget entries are all constrained to zero, for $n = h, \ell$.

Table 10: Estimation of the spatial equilibrium model: equations (10) through (13)

	Pooled generations, without budget entries		Split generations, with budget entries	
	(1)	(2)	(3)	(4)
θ_h: high-to-high spillovers	0.1144***	0.0755**	0.1143***	0.0832***
	(0.0279)	(0.0265)	(0.0271)	(0.0245)
φ_h: log-TFP advantage of SCs, high		0.0877**		0.0743*
		(0.0336)		(0.0296)
θ_ℓ: high-to-low spillovers	0.1042***	0.0396	0.1098***	0.0473
	(0.0250)	(0.0403)	(0.0255)	(0.0410)
φ_ℓ: log-TFP advantage of SCs, low		0.0721		0.0550
		(0.0527)		(0.0452)
ζ_{ph}: education budget entry, high			0.0294***	0.0250**
			(0.0064)	(0.0079)
$\zeta_{p\ell}$: education budget entry, low			0.0278*	0.0356**
			(0.0125)	(0.0137)
ρ_h: persistence, high (baseline value)	0.6325***	0.4553**	0.6306***	0.5218***
	(0.1013)	(0.1566)	(0.1030)	(0.1121)
$\Delta\rho_h$: persistence, high (young minus old)			-0.0388	-0.0523*
			(0.0235)	(0.0261)
ρ_ℓ: persistence, low (baseline value)	0.3298***	0.1769	0.3078***	0.2594**
	(0.0855)	(0.1325)	(0.0847)	(0.0992)
$\Delta\rho_\ell$: persistence, low (young minus old)			0.0332	0.0351
			(0.0199)	(0.0200)
F-statistic: budget entries δ_{ph}, ζ_{ph} (p-value)	N/A	N/A	(0.000)	(0.000)
F-statistic: budget entries $\delta_{p\ell}, \zeta_{p\ell}$ (p-value)	N/A	N/A	(0.000)	(0.000)
No. of observations, eq.s (10), (11) and (12)	64	64	64	64
No. of observations, eq. (13)	45	45	45	45

Notes: *, ** and *** denote significance at the 10, 5, and 1 per cent level, respectively. The Table reports several estimates of the spatial equilibrium model, for a fixed values $\mu = 0.6$. The estimation is performed via Iterated GMM, allowing arbitrary correlation across all the model's error terms, clustering observation sharing the same control municipality. The estimates from column 1 and 2 do not include the budget entries b_{pc} on the right-hand side of the model's equations, constraining the associated coefficients to zero. The estimates in columns 3 and 4 are based on duplicate equations for the demographic outcomes, split by age group (young vs. old); the table reports the differences $\Delta\rho_h$ and $\Delta\rho_\ell$ in the group-level persistence parameters (we treat the old group as the baseline). The estimates in columns 1 and 3 constrain the extra effect of Science Cities on TFP (φ_h, φ_ℓ) to zero. The difference in high-skilled population, $(h_s - h_z)$, is not used as an instrument in the equation for low-skilled population, (11). The equation for the low-skilled wage, (13) is estimated on a subset of city pairs (45).

This allows us to focus on our keys parameters θ_n, and ρ_n expressing the agglomeration and persistence forces. We find that both spillover parameters are estimated in a neighborhood of $\hat{\theta}_n \simeq 0.11$, which represents the extra elasticity of local TFP from high-skilled population; this number lies in the range of agglomeration parameters typically found in middle-income countries (Duranton, 2015). The two persistence parameters are different between the two skill groups: we estimate $\hat{\rho}_h \simeq 0.63$, indicating that for typical workers in the highly skill group, living in a Science City bears an amenity value approximately equal to 63 per cent of their salary. This is a large while realistic figure. The associated estimate for the low skilled group, $\hat{\rho}_\ell \simeq 0.33$, is about half that value. All these estimates are statistically significant at the 1 per cent level.

The estimates from column 2 expand on those from column 1 by allowing for the TFP effects of Science Cities expressed by the φ_n parameters. While the estimated effect is not statistically significant for the low skilled group, we estimate an 8 per cent premium associated with the salary in the R&D-IT sector (which is significant at the 5 per cent level). Note that this extra effect is estimated separately from agglomeration forces; still, it is of large magnitude. The high-to-high spillovers are now estimated at a lower value $\hat{\theta}_h \simeq 0.08$, which is significant at the 5 per cent level; the estimate for the high-to-low spillovers $\hat{\theta}_\ell$ is no longer statistically significant. These results suggests that the tradition of Soviet Science cities maintains to this days in the form of intangible productivity advantages for R&D, less so for other sectors. It must also be remarked that upon allowing for these extra effects, all our key persistence parameters ρ_n are (as expected) decreased in magnitude, while no longer statistically significant for the low-skilled group.

Columns 3 and 4 of Table 10 report estimates that expand on those from columns 1 and 2 respectively, by including the municipal budget entries b_{pc} into the model and by duplicating the demographic equations by age group (bringing to six the total number of equations in the model). The municipal budget entries are jointly strong predictors of our key outcomes, as evidenced by the F-statistic p-values. In particular, in the table we highlight the estimates of ζ_h, ζ_ℓ for the education budget entry: they are statistically significant, and they indicate that a marginal increase in local expenditures in education are associated with a 3 per cent increase in salaries for both skill groups. In columns 3 and 4, the key parameters θ_n and φ_n from the salary equations are estimated similarly as in columns 1 and 2, respectively. In our favorite specification from column 4, $\hat{\theta}_h \simeq 0.08$ is significant at the 1 per cent level, $\hat{\varphi}_h \simeq 0.07$ is significant at the 10 per cent level, while the parameters for the low skilled are not statistically significant.

Duplicating the demographic equations by skill group appears to be of little consequence towards our estimates of ρ_n: no large differences between older and younger generations are attested, suggesting that the advantages of Science Cities are not bound to deteriorate rapidly. In our favorite specification from column 4, we estimate $\hat{\rho}_h \simeq 0.52$ for the older highly skilled and $\hat{\rho}_\ell \simeq 0.26$ for the older low skilled, which are statistically significant at the 1 and 5 per cent level respectively; the only difference between generations that is statistically significant (at the 10 per cent level) corresponds with a lower estimate for the younger highly skilled, $\hat{\rho}_h \simeq 0.47$: a difference which is not too large.

Together, these results tell a coherent story. On the one hand, there are certainly different factors that affect the location choices of workers, such as for example amenities and variations in the supply of various public goods. However, there is substantive evidence that the initial allocation of highly-skilled workers in Science Cities, for various reasons, has produced substantial path-dependence – what we call "persistence forces." Furthermore, the higher concentration of highly-skilled workers appears to generate agglomeration economies, ultimately resulting in higher wages. Our interpretation of the recent history of Science Cities in transition years overlaps the one by Rowland (1996) about closed cities, and is as follows. Unable to move due to legal restrictions, or otherwise unwilling to do so, displaced and impoverished scientists, researchers and technicians had to adapt to a new market economy which was rapidly shifting from manufacturing to services. Thanks to their higher human capital, their hometowns attracted outside firms; sometimes, these workers created new businesses. Because of self-sustaining agglomeration economies, this dynamics eventually spurred faster local development.

5 Spillovers at the firm level

In this section we analyze the effect of Science Cities on firms' innovation and performance. In particular, we examine – in a descriptive sense – to what extent locating closer to a Science City predicts differentials in these outcomes. Like section 3, this one is split into three parts: methodology, data and results.

5.1 Methodology

There are many versions of distance or proximity measures in the literature. Because we are interested in how the proximity to Science Cities (and non-Science Cities) affects innovation and performance of firms, we focus on measures of geographic spillovers. We assume that the effect of distance decays exponentially and define the agglomeration potential measures as:

$$G_{cfr} \equiv G_{cfr}\left(H_1, \ldots, H_c; \lambda\right) = \sum_{c=1}^{C} \exp\left[-\lambda \cdot \operatorname{dist}\left(f, c\right)\right] H_c, \tag{14}$$

where $c \in s, z$ denotes the type of location (as usual, subscript s refers to Science City while z to a non-Science City); $f = 1, \ldots, F$ indexes firms; r is a subscript for Russian regions; λ is a decay parameter; $\text{dist}(f, c)$ is the geodesic distance between firm f and location s; H_c is a relevant characteristic of location c. We examine four different characteristics H_c that likely relate to innovation potential. These are: the fractional patents produced in location c, the graduate and postgraduate share of its population and its share of R&D employment.

Functional forms that involve analogous terms are routinely adopted in studies of R&D spillovers (Lychagin et al., 2016) as well as of agglomeration effects between firms (Drucker, 2012). Note that these measures are flexible and they vary with the choice of H_c and with decay parameter λ which is assumed to be greater than zero, such that H_c in location c has less and less potential influence on a firm f as the distance between firm f and location c increases. The larger the value of λ, the more rapidly the potential effect of H_c diminishes with $\text{dist}(f, c)$. When $\lambda = 3$, the weight of H_c in equation(14) is $6 \cdot 10^{-6}$ if firm f is 4 kilometers away from location c; when $\lambda = 5$, the same weight is achieved if firm f is 2.4 kilometers away from location c.[39] To make sure that what we are measuring is indeed the spillover effect of Science Cities and not some other factor, we also control for potential spillover effects from non-Science Cities and calculate a comparable agglomeration potential measure G_{zfr}. For reasons discussed in sections 3.2.1 and 3.3.1, we exclude academic towns and closed cities from the calculation of agglomeration potential measures. To further ensure that we are not measuring standard patterns around large cities, $W_{fr,d}$ includes an indicator for whether the firm is located in a city with population over 1 million.

Our innovation outcomes of interest are all binary. Thus, we estimate a number of probit models with the following latent variable representation:

$$I_{fr}^{*} = \beta_0 + \sum_{d=1}^{D} \beta_d W_{fr,d} + \gamma_s G_{sfr} + \gamma_z G_{zfr} + \eta_r + \varepsilon_{fr} \tag{15}$$

where $f = 1, \ldots, F$ indexes firms; $s = 1, \ldots, S$ denotes Science Cities; $z = 1, \ldots, Z$ denotes non-Science Cities; r is a subscript for Russian regions; I_{fr}^{*} is the latent variable associated with one specific innovation binary outcome I_{fr}; $(W_{fr,1}, \ldots, W_{fr,D})$ are D controls

[39]We report estimates using $\lambda = 3$; the findings are robust to using $\lambda = 2$, $\lambda = 4$ or $\lambda = 5$ – see section 5.3.3. Descriptive statistics and cross-correlations for the resulting firm-level agglomeration measures can be found in Tables F.1 and F.2 in Appendix F.

available in the data (see section 5.2); η_r is a region fixed effect; and lastly ε_{fr} is an error term which follows a standard normal distribution. Our firm performance outcomes are continuous, thus we run OLS regressions:

$$\log P_{fr} = \tilde{\beta}_0 + \sum_{d=1}^{D} \tilde{\beta}_d W_{fr,d} + \tilde{\gamma}_s G_{sfr} + \tilde{\gamma}_z G_{zfr} + \tilde{\eta}_r + \upsilon_{fr} \tag{16}$$

analogous to (15). In model (16), P_{fr} represents performance indicators such as the firm's operating revenue (sales) or labor productivity.

For linear and non-linear models alike, our main explanatory variables of interest are the agglomeration potential measures of Science and non-Science Cities, G_{sfr} and G_{zfr}. We report estimations using fractional patents and postgraduate share-based measures in the same regression.[40] We also examine heterogeneity in the effects of spillovers by sector and age of firms. Different sources of spillovers can affect firms in different ways. Sinani and Meyer (2004) show that spillovers of technology transfer from FDI are influenced by the recipient firm's size, ownership structure and trade orientation. In the more specific setting of transition economies, Gorodnichenko et al. (2014) also find differences in spillover effects by firm size, sector and age.

While we do not attempt to give any causal interpretation to the firm-level estimates, we observe that the concerns of endogeneity are limited in this setting. Since the creation of Science Cities predates the establishment of most modern Russian firms – virtually all in our sample – the only way for the distance-based regressor and the error term to be correlated is if a Science City attracts or otherwise encourages the location of more innovative or better performing firms in their proximities. We make no attempts to correct for this possible instance of endogeneity, since we are interested in evaluating in a descriptive sense whether any relationship between Science Cities and firm-level outcomes extends in space. We do not intend to remove a potential mechanism by which such relationships may manifest themselves.

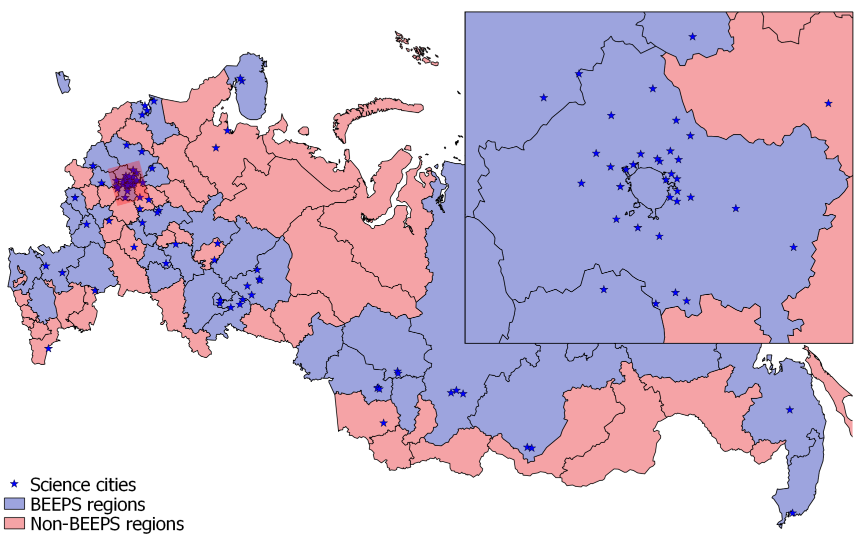
5.2 Data and descriptive statistics

We use the fifth round of BEEPS for Russia only. BEEPS is a firm-level survey conducted by the European Bank for Reconstruction and Development and the World Bank, based

[40] We do not include all of them at the same time because they are highly correlated within location type c (see Table F.2 in Appendix F). Results using one of the other measures in combination with fractional patents are available on request.

on face-to-face interviews with managers of registered firms with at least five employees. Stratified random sampling is used to select eligible firms to participate in the survey. In Russia, 4,220 interviews were completed in a subset of Russian regions; the chosen regions encompass the majority of historical Science Cities, as shown in Figure 4. The database contains geographic coordinates of the firm's location, based on which we can determine distances from Science Cities. Additional information about BEEPS V Russia is given in Appendix G.

Figure 4: Location of Science Cities and regions covered in BEEPS V Russia



Sources: Table B.1 and BEEPS V Russia.

Outcomes. BEEPS V included, for the first time, an innovation module, which provides information on whether, in the last three years prior to the survey, a firm engaged in R&D (in-house or outsourced), introduced a new product, process or technological innovation, and on whether it was ever granted a patent. We manually clean the information contained in the innovation module: for each firm, we verify whether survey responses match the firm's main product and industry by using external information about each firm and comparing the descriptions of the main new product or process reported in the survey with the definitions given in the Oslo Manual (OECD and Statistical Office of the European Communities, 2005). In addition, BEEPS V includes information on sales and employment.

Controls. BEEPS V Russia also contains measures for several firm characteristics, such as: age; industry; exporter status; ownership; geographical scope of the main market; the number of permanent, full-time employees; as well as the share of employees with a university degree. We also control for whether the firm is credit-constrained and whether it is located in a large city (with population over 1 million).

Summary statistics. Table 11 reports descriptive statistics at the firm level, taking into account survey weights. Notably, almost half of all firms (47.1 per cent) report introducing a new product or a new process in the last three years prior to the survey; the fraction of firms performing R&D is lower (31.5 per cent).

Table 11: Firm-level data: Descriptive statistics

	Obs.	Mean	Linearised std. error	[95% Conf. interval]	
Young firms (0-5 years)	4,220	0.169	0.054	0.063	0.274
25%+ foreign owned	4,220	0.058	0.040	-0.020	0.136
25%+ state owned	4,220	0.009	0.007	-0.005	0.022
Exporter	4,220	0.209	0.056	0.098	0.320
Main market: local	4,220	0.502	0.043	0.418	0.587
Main market: national	4,220	0.495	0.043	0.410	0.579
% of employees with a university degree	4,045	55.639	3.793	48.181	63.097
Located in a city with population over 1 million	4,220	0.605	0.011	0.583	0.626
Credit-constrained firm	4,220	0.412	0.060	0.294	0.529
Log (permanent, full-time employees)	4,211	3.528	0.167	3.200	3.856
Log (sales)	3,027	17.889	0.209	17.478	18.299
Log (labor productivity)	3,021	14.346	0.182	13.989	14.704
R&D (dummy)	4,220	0.315	0.058	0.201	0.429
Technological innovation (dummy)	4,220	0.471	0.058	0.356	0.586
Product innovation (dummy)	4,220	0.326	0.058	0.211	0.441
Process innovation (dummy)	4,220	0.306	0.053	0.201	0.410
Ever granted a patent (dummy)	1,998	0.163	0.053	0.059	0.267

Notes: Survey-weighted observations. Linearised Taylor standard errors clustered on strata.

5.3 Empirical results

Next, we present the estimates of regression models (15) and (16) with $\lambda = 3$; we obtain similar results with $\lambda = 2$ and higher values for this parameter (see section 5.3.3).

5.3.1 Innovation outcomes

Table 12 presents the results from the estimation of several probit models with latent variable representation (15) for four separate firm-level binary outcomes I_{fr}: whether a firm engages in any R&D activity (in-house or contracted); whether in the three years prior to the survey the firm has introduced a new product or process; and lastly, if the firm's innovation effort has ever resulted in being granted a patent. On the right-hand side of (15), we employ the agglomeration measures discussed in section 5.1. In the table, we present the average probit marginal effects, which are interpreted as the average increase in the probability of $I_{fr} = 1$ associated with a unit increase in the G_{sfr} (Science Cities – SC) and G_{zfr} (non-Science Cities – non-SC) measures.

Panel A shows that the estimates of Science City fractional patents γ_s are positive and statistically significant for three outcome variables in the total sample: engagement in R&D, product innovation and having been granted a patent. A doubling of the index of fractional patents is associated with on average 2.4 per cent increase in the probability that the firm engages in R&D, 2.2 per cent increase in the probability that it has introduced a new product in the last three years and 2.3 per cent increase in the probability that it has ever been granted a patent. Notably, the estimate of non-Science City fractional patents γ_z is positive and statistically significant only for having been granted a patent, but its magnitude is about six times lower than the magnitude of the estimate of Science City fractional patents γ_s. The estimates of Science City postgraduate share γ_s are not significant. For non-Science Cities, the estimate of postgraduate share γ_z is negative and marginally significant (at 10 per cent level) for process innovation. These findings indicate that the innovativeness of Science Cities spills over to the firms that are located sufficiently close to them, and that we are not just capturing standard patterns around any larger cities. While these marginal effects cannot be interpreted in a causal sense, they are indicative of some economic mechanisms that induce firms with more innovation potential to locate in the proximity of Science Cities.

Panels B-D explore whether there are any differences in the spillover effects by sector, technological and knowledge intensity (using the OECD definitions) and age of firms, respectively. The estimates in Panel B suggest that only manufacturing firms benefit from the Science Cities spillover effects, with the patent-based agglomeration potential measure positively and statistically significantly associated with R&D and ever being granted a patent. The non-Science Cities patent-based agglomeration potential measure is positively and statistically significantly associated with having a patent for service sector firms, though the magnitude of the coefficient is smaller than that for its Science Cities counterpart. The γ_s and γ_z coefficient estimates are not statistically significant for skills-based agglomeration potential measure.

Table 12: Firm-level innovation outcomes: Probit average marginal effects ($\lambda = 3$)

Agglomeration potential measure	R&D	Product innovation	Process innovation	Has a patent
Panel A: Total sample				
Fractional patents, SC	0.0244***	0.0220***	-0.0034	0.0229***
	(0.0045)	(0.0068)	(0.0065)	(0.0067)
Fractional patents, non-SC	0.0005	0.0002	0.0008	0.0037***
	(0.0006)	(0.0008)	(0.0007)	(0.0005)
Postgraduate share (%), SC	-0.2253	-0.2202	-0.1121	-0.1185
	(0.1842)	(0.2119)	(0.2857)	(0.1438)
Postgraduate share (%), non-SC	0.0151	0.2505	-0.6433*	-0.2728
	(0.3302)	(0.1909)	(0.3747)	(0.6731)
Panel B: Allowing different coefficients by sector				
Fractional patents * Manufacturing, SC	0.0330***	0.1039	-0.0007	0.0282**
	(0.0104)	(0.0927)	(0.0115)	(0.0132)
Fractional patents * Services, SC	-0.1063	0.0680	0.0246	0.0244
	(0.0756)	(0.0488)	(0.0190)	(0.0254)
Fractional patents * Manufacturing, non-SC	-0.0022	-0.0007	-0.0431	-0.0140
	(0.0025)	(0.0014)	(0.0397)	(0.0208)
Fractional patents * Services, non-SC	0.0004	0.0001	0.0008	0.0044***
	(0.0006)	(0.0008)	(0.0008)	(0.0006)
Postgraduate share (%) * Manufacturing, SC	-0.1826	-0.4702	-0.8137	0.0508
	(0.2325)	(0.6045)	(0.8850)	(0.1767)
Postgraduate share (%) * Services, SC	0.4116	-3.7198	-0.1852	-2.0958
	(0.5533)	(3.0360)	(0.3027)	(1.3763)
Postgraduate share (%) * Manufacturing, non-SC	-0.1465	0.2325	-0.1502	0.2381
	(0.5479)	(0.4590)	(0.5392)	(0.9873)
Postgraduate share (%) * Services, non-SC	0.0534	0.2614	-0.5509	-0.8380
	(0.3992)	(0.1991)	(0.4955)	(0.6444)
Panel C: Allowing different coefficients by technological level				
Fractional patents * High-tech, SC	-0.0377	0.1703	0.1773**	-0.0163
	(0.0989)	(0.1456)	(0.0865)	(0.1117)
Fractional patents * Other, SC	0.0269***	0.0599**	-0.0050	0.0448**
	(0.0058)	(0.0266)	(0.0066)	(0.0189)
Fractional patents * High-tech, non-SC	-0.0226	-0.0206	-0.4301***	-0.0524
	(0.0315)	(0.0286)	(0.1512)	(0.0624)
Fractional patents * Other, non-SC	0.0006	0.0003	0.0010	0.0039***

Table 12 – continued from previous page				
Agglomeration potential measure	R&D	Product innovation	Process innovation	Has a patent
	(0.0006)	(0.0008)	(0.0007)	(0.0006)
Postgraduate share (%) * High-tech, SC	0.2577	-0.6694	-2.5992**	0.3650
	(0.5993)	(1.0426)	(1.1753)	(0.6447)
Postgraduate share (%) * Other, SC	-0.3350	-2.7117	0.0073	-1.6496
	(0.2652)	(1.7410)	(0.2844)	(1.2145)
Postgraduate share (%) * High-tech, non-SC	0.3244	0.2215	4.4324***	0.4501
	(0.6208)	(0.5379)	(1.6035)	(0.9612)
Postgraduate share (%) * Other, non-SC	-0.0055	0.2899	-0.7400*	-0.2484
	(0.3818)	(0.2095)	(0.4035)	(0.7021)
Panel D: Allowing different coefficients by firm age				
Fractional patents * Young, SC	-0.0309	-0.0570		0.0783**
	(0.0485)	(0.0545)		(0.0346)
Fractional patents * Old, SC	0.0230***	0.0211**		0.0192***
	(0.0049)	(0.0087)		(0.0059)
Fractional patents * Young, non-SC	-0.0030	-0.2384**		-0.2280
	(0.0098)	(0.1111)		(0.1642)
Fractional patents * Old, non-SC	0.0004	0.0000		0.0042***
	(0.0006)	(0.0008)		(0.0006)
Postgraduate share (%) * Young, SC	-8.5290	0.0392		-0.5714
	(6.7180)	(0.7411)		(0.5490)
Postgraduate share (%) * Old, SC	-0.1855	-0.1703		-0.0399
	(0.1966)	(0.1985)		(0.1455)
Postgraduate share (%) * Young, non-SC	0.1281	0.2971		2.1102*
	(0.3814)	(0.4863)		(1.1668)
Postgraduate share (%) * Old, non-SC	-0.1156	0.6180**		-1.8553*
	(0.3798)	(0.2821)		(0.9982)
Number of observations	4,040	4,040	4,040	1,863
Number of strata	1,224	1,224	1,224	896

Notes: *, ** and *** denote significance at the 10, 5, and 1 per cent level, respectively. Linearised Taylor standard errors clustered on strata are reported in parentheses. Average marginal effects based on probit using survey-weighted observations. Only coefficients on agglomeration potential measures are reported. Fractional patents agglomeration potential measure is based on the number of patent applications to EPO in 2006-15 in municipalities with Science Cities, by inventor (fractional counting). Postgraduate education agglomeration potential measure is based on the percentage of population with postgraduate education in municipalities with Science Cities in 2010. All regressions include region and sector fixed effects and control for other firm characteristics: log number of permanent, full-time employees, % of employees with a completed college degree, and indicators for young firms (up to five years old), 25% foreign and state ownership, exporter status, local and national main markets for the firms' products, credit constrainedness and whether a firm is located in a city with population over 1 million. SC – Science Cities; non-SC – non-Science Cities; Mnf. – manufacturing; Serv. - services.

For high-tech sector firms (defined as those firms in high-tech, medium-high-tech and high knowledge intensity sectors), process innovation is positively and significantly associated with the Science Cities patent-based agglomeration potential measure, but negatively and significantly associated with its skills-based counterpart. Their other innovation-related activities do not benefit from being Science Cities spillovers; this could be explained by the relatively small number of high-tech sector firms in our sample overall, and in Science Cities in particular. In contrast, the Science Cities patents-based agglomeration potential measure is positively and significantly associated with R&D, product innovation and ever been granted a patent for firms in other sectors, which do not benefit from skills-based spillovers from Science Cities. High-tech sector firms benefit from skills spillovers but experience negative patents spillovers non-Science Cities when it comes to process innovation. The non-Science Cities agglomeration potential measures are generally not significantly associated with most outcomes for firms in other sectors; the exceptions are patents-based measure's positive and significant association with having a patent, and skills-based measure's negative and significant (at 10 per cent) association with process innovation.

Furthermore, the estimates in Panel D indicate that old firms (more than 5 years old) benefit from the spillovers more than young firms (for which most estimated coefficients are not statistically significant), with one exception: young firms located close to Science Cities experience stronger spillovers for having a patent than old firms. A doubling of the index of fractional patents is associated with 7.8 per cent increase in the probability that a young firm has ever been granted a patent, while the same number is only 1.9 per cent for old firms. Young firms located close to non-Science Cities do not benefit from similar patent-based spillovers - if anything, they might actually experience negative spillovers. They do, however, benefit from skills-based spillovers. Old firms close to non-Science Cities benefit from patents-based spillovers as well, but the magnitude of the effect is about 4.6-times smaller than near Science Cities - a doubling of the index of fractional patents is associated with 0.4 per cent increase in the probability that an old firm has ever been granted a patent.

To sum up, estimates in Table 12 suggest that Science Cities have spillover effects on the innovation activity of firms, particularly R&D and having a patent, and that this spillover is mostly driven by the agglomeration effects of patents, rather than skills (although the two measures are highly correlated, as shown in Table F.2). There are some spillovers from non-Science Cities, too, though only for particular types of firms. Moreover, they are not necessarily positive, and if the spillovers are found for both Science and non-Science Cities, their magnitude is larger in the former.

5.3.2 Performance indicators

The measurement of the returns to R&D and innovation corresponds with a traditional line of research in empirical studies about innovation and productivity; see, for instance, two distinct surveys: Hall et al. (2010) and Syverson (2011). In our setting, we are similarly interested in uncovering performance advantages for firms that are located close to Science Cities, which can be due either to the indirect effect of firm-level innovation spurred by Science Cities (which we illustrated above) or to spillovers of a different kind. To this end, we provide reduced form evidence about the association between Science Cities and firms' labor productivity or sales, by estimating model (16) under different specifications. The results are reported in Table 13.

The estimates of $\tilde{\gamma}_s$ are positive and statistically significant for the patent-based agglomeration potential measure only close to Science Cities. The estimates for spillovers from non-Science Cities are close to zero and not statistically significant. However, for the skills-based agglomeration potential measure, the estimates for spillovers from Science Cities are negative and significant at 10 per cent.

The estimates in Panel B suggest that manufacturing firms benefit from the Science Cities spillover effects more than services firms, with both sales and labor productivity positively and statistically significantly associated with the patents-based agglomeration potential measure, though the magnitude of the non-significant estimate is around 3.5-times higher for services firms. A doubling of the index of fractional patents is associated with a 4.0 per cent increase in sales and labor productivity of manufacturing sector firms located close to Science Cities. Skills-based agglomeration potential measure is negatively and statistically significantly associated with sales and labor productivity of manufacturing firms near Science Cities.

A similar effect of patents-based spillovers from Science Cities can be observed for firms in medium- and low-tech sectors (Panel C). Interestingly, performance of firms in high-tech sectors is negatively and statistically significantly (at 10 per cent level) associated with the patents-based agglomeration potential measure. High-tech firms tend to operate closer to the technological frontier, and may still need to find a market or sufficient number of customers for their products, and being close to Science Cities, where competition among firms in high-tech sectors is fiercer, might make it more difficult for them to monetise their own patents.

Table 13: Firm-level performance outcomes: OLS ($\lambda = 3$)

Agglomeration potential measure	Sales	Labor productivity
Panel A: Total sample		
Fractional patents, SC	0.0522***	0.0526***
	(0.0168)	(0.0172)
Fractional patents, non-SC	-0.0039	-0.0035
	(0.0027)	(0.0026)
Postgraduate share (%), SC	-1.4287*	-1.3979*
	(0.8127)	(0.8269)
Postgraduate share (%), non-SC	-0.0373	-0.0083
	(1.2887)	(1.3204)
Panel B: Allowing different coefficients by sector		
Fractional patents * Manufacturing, SC	0.0396**	0.0389**
	(0.0162)	(0.0164)
Fractional patents * Services, SC	0.1322	0.1371
	(0.0910)	(0.0964)
Fractional patents * Manufacturing, non-SC	-0.0113*	-0.0113*
	(0.0064)	(0.0063)
Fractional patents * Services, non-SC	-0.0032	-0.0028
	(0.0029)	(0.0027)
Postgraduate share (%) * Manufacturing, SC	-2.0802**	-2.1156**
	(0.8904)	(0.9022)
Postgraduate share (%) * Services, SC	-1.7189	-1.6926
	(1.0924)	(1.1333)
Postgraduate share (%) * Manufacturing, non-SC	-0.1146	-0.2592
	(1.4142)	(1.4247)
Postgraduate share (%) * Services, non-SC	0.0993	0.2040
	(1.6037)	(1.6351)
Panel C: Allowing different coefficients by technological level		
Fractional patents * High-tech, SC	-0.3985*	-0.4122*
	(0.2261)	(0.2301)
Fractional patents * Other, SC	0.0540***	0.0545***
	(0.0188)	(0.0193)
Fractional patents * High-tech, non-SC	0.1061	0.1107
	(0.2097)	(0.2119)

Table 13 – continued from previous page		
Agglomeration potential measure	Sales	Labor productivity
Fractional patents * Other, non-SC	-0.0038	-0.0034
	(0.0027)	(0.0025)
Postgraduate share (%) * High-tech, SC	0.6037	0.6584
	(1.8894)	(1.9327)
Postgraduate share (%) * Other, SC	-1.2644	-1.2214
	(1.0333)	(1.0589)
Postgraduate share (%) * High-tech, non-SC	-0.2717	-0.4324
	(3.2760)	(3.2877)
Postgraduate share (%) * Other, non-SC	-0.1910	-0.1690
	(1.4361)	(1.4722)
Panel D: Allowing different coefficients by firm age		
Fractional patents * Young, SC	0.5255*	0.5913*
	(0.3167)	(0.3182)
Fractional patents * Old, SC	0.0463**	0.0469**
	(0.0232)	(0.0232)
Fractional patents * Young, non-SC	0.0399*	0.0422*
	(0.0228)	(0.0251)
Fractional patents * Old, non-SC	-0.0049*	-0.0048*
	(0.0027)	(0.0026)
Postgraduate share (%) * Young, SC	-4.8182**	-4.8203**
	(1.9409)	(1.9702)
Postgraduate share (%) * Old, SC	-1.3441	-1.4116
	(1.3289)	(1.3260)
Postgraduate share (%) * Young, non-SC	-0.3411	-0.3831
	(1.9866)	(2.1699)
Postgraduate share (%) * Old, non-SC	-0.3225	-0.3340
	(1.6989)	(1.7117)
Number of observations	2,926	2,926
Number of strata	1,074	1,074

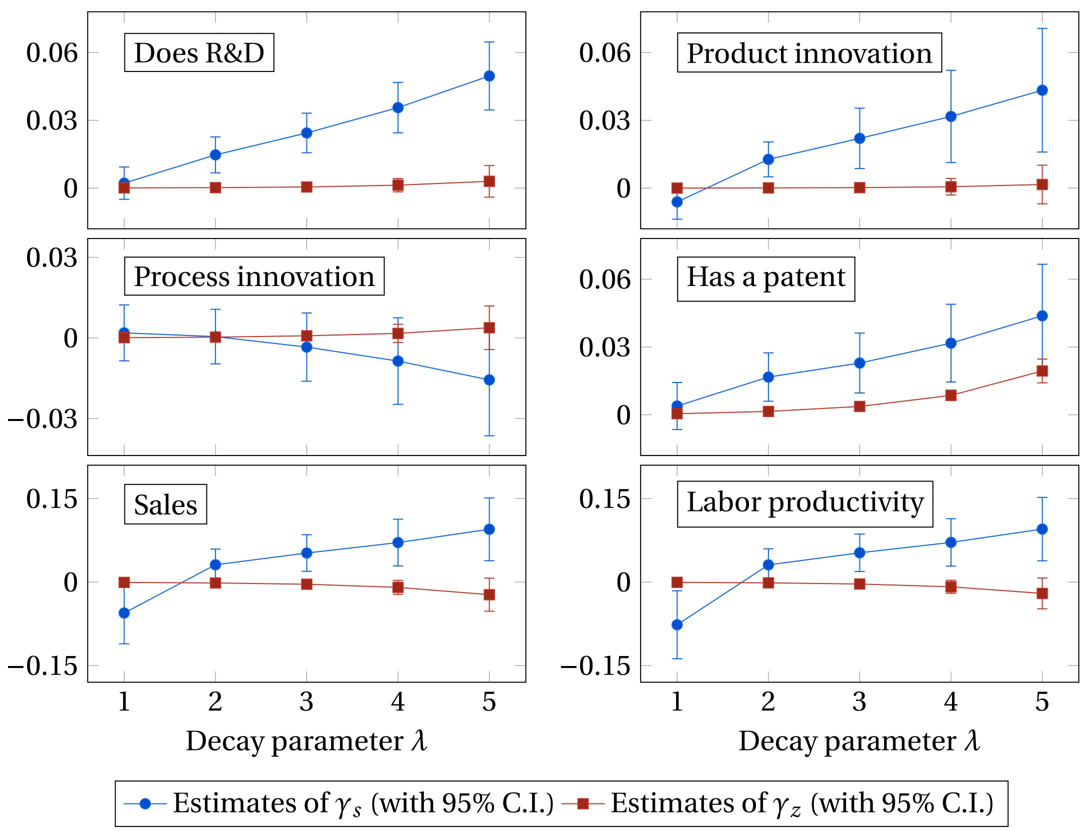
Notes: Simple OLS using survey-weighted observations. For other details, see the note accompanying Table 12. SC – Science Cities; non-SC - non-Science Cities; Mnf. – manufacturing; Serv. – services.

Young firms located near Science Cities benefited from patent-based spillovers not only in terms of having their own patents, but also in terms of sales and labor productivity (Panel D). Young firms located near non-Science Cities also benefited from being close to better infrastructure and larger markets, but the magnitude of the estimates there was less than a tenth of those near Science Cities. They were also significantly larger than the magnitudes of the estimates for old firms. On the other hand, young firms near Science Cities experienced negative skill-based spillovers on both sales and labor productivity.

5.3.3 Alternative decay parameters

The spillover estimates presented above are based on $\lambda = 3$. To evaluate the dependence of our results on this parameter choice, we re-estimate the model for the alternative values $\lambda = 1, 2, 3, 4$ and 5. The results relative to our patent-based agglomeration measures, for both Science and non-Science Cities, are shown in Figure 5. The higher the rate of decay (the closer the firms are to Science Cities), the larger the spillover effects; with the exception of process innovation as an outcome, they are statistically significant and positive. Spillovers from non-Science Cities are positive and statistically significant only for having a patent, but they are much smaller than spillovers from Science Cities.

Figure 5: Estimates of γ_s and γ_z for alternative values of the decay parameter λ



Sources: Panel A of Tables 12 and 13 and authors' calculations.

5.3.4 Discussion

The results in the previous two sections show that Science Cities have spillover effects on innovation and performance outcomes of firms located close to them, while this is not generally the case for non-Science Cities. For innovation outcomes, such as R&D, having a patent and product innovation, and for performance outcomes, such as sales and labor productivity, the source of these spillovers is knowledge accumulated in patents produced in Science Cities. As expected, spillover effects are not the same for all types of firms, and for some firms, they are negative. For example, performance outcomes of firms in high-tech sectors are negatively and statistically significantly associated with the patents-based agglomeration potential measure. This may indicate that they are very close to or at the technological frontier and the market for their products or services is not yet sufficiently developed. It may also reflect fiercer competition among high-tech firms in the vicinity of Science Cities. Medium- and low-tech sector firms, on the other hand, experience positive spillovers from Science Cities on their performance, primarily through the skills channel.

6 Conclusion

In this article we have analyzed the long-run effects of a unique historical placed-based policy: the creation of R&D-focused Science Cities in Soviet Russia. Both the initial establishment and the eventual suspension of this program was largely guided by political factors that are arguably exogenous to drivers of current social and economic conditions of Russian cities. We compare Science Cities with other localities that were observationally similar to them at the time of their selection, and we compute differences in the current characteristics between the two groups. We find that former Science Cities are bigger today, largely because they host a higher number of well-educated individuals. Moreover, they produce a higher number of internationally recognized patents (both in absolute terms and considering the average in the population of potential inventors); their R&D and ICT sectors are more developed, and pay higher salaries. Lastly, Science Cities host more productive small businesses in the services sector. Through a separate firm-level analysis, we attest some evidence in support of the hypothesis that the effect of Science Cities extends beyond their municipal borders.

Because our results hold largely unchanged after the removal, from the estimation sample, of Science Cities that receive resumed support from the Russian government at present, we conjecture that they are consequent to the interaction between persistence (path-dependence) and agglomeration forces. In more mundane terms, highly skilled individuals who have remained in their former cities of residence have contributed to the emergence of more productive businesses in the new market economy. This insight is substantiated by estimating a simple model of spatial equilibrium; while this is designed to accommodate alternative explanations, the results lend them little support. By analyzing municipal budgets, we rule out alternative explanations that have to do with differential governmental transfers. In addition, by examining our data in more detail we find little support for the hypothesis of rapid mean reversion of socio-economic outcomes to a more symmetric equilibrium.

Our contribution extends previous findings about long-run effects of place-based policies to a unique historical program that focused on human capital and R&D. More generally, our results are also informative for science and innovation policy, both in the context of emerging economies such as Russia and in those of traditionally capitalist countries. We hope that these results will be invoked to motivate similar R&D policies but with a civil, instead of military, purpose.

References

Abadie, Alberto and Guido W. Imbens (2006) "Large sample properties of matching estimators for average treatment effects," *Econometrica*, Vol. 74, No. 1, pp. 235–267.

——— (2011) "Bias-corrected matching estimators for average treatment effects," *Journal of Business and Economic Statistics*, Vol. 29, No. 1, pp. 1–11.

Agirrechu, Aleksandr A. (2009) *Russian science cities: History of formation and development (Naukogradi Rossiyi: Istoriya formirovaniya i razvitiya)*, Moscow: Moscow University Press, Available in Russian only.

Ambler, John, Denis J. B. Shaw, and Leslie Symons eds. (1985) *Soviet and East European transport problems*, London: Croom Helm.

Andrienko, Yuri and Sergei Guriev (2004) "Determinants of interregional mobility in Russia: Evidence from panel data," *Economics of Transition*, Vol. 12, No. 1, pp. 1–27.

Audretsch, David B. and Maryann P. Feldman (1996) "R&D spillovers and the geography of innovation and production," *American Economic Review*, Vol. 86, No. 3, pp. 630–640.

Beaudry, Catherine and Andrea Schiffauerova (2009) "Who's right, Marshall or Jacobs? The localization versus urbanization debate," *Research Policy*, Vol. 38, No. 2, pp. 318–337.

Ben-Michael, Eli, Jesse Rothstein, and Avi Feller (2019) "The Augmented Synthetic Control Method," mimeo.

Bircan, Çağatay and Ralph De Haas (2020) "The limits of lending? Banks and technology adoption across Russia," *The Review of Financial Studies*, Vol. 33, No. 2, pp. 536–609.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen (2013) "Identifying technology spillovers and product market rivalry," *Econometrica*, Vol. 81, No. 4, pp. 1347–1393.

Boschma, Ron and Koen Frenken (2011) "The emerging empirics of evolutionary economic geography," *Journal of Economic Geography*, Vol. 11, No. 2, pp. 295–307.

Bronzini, Raffaello and Guido de Blasio (2006) "Evaluating the impact of investment incentives: The case of Italy's Law 488/1992," *Journal of Urban Economics*, Vol. 60, No. 2, pp. 327–349.

Busso, Matias, Jesse Gregory, and Patrick M. Kline (2013) "Assessing the incidence and efficiency of a prominent place based policy," *American Economic Review*, Vol. 103, No. 2, pp. 897–951.

Central management unit of the military communications of the Red Army (1943) *Schemes of railways and waterways of the USSR (Shemi zheleznih dorog i vodnih putey soobsheniya SSSR)*: Military Publishing of the People's Commisariat of Defence, Available in Russian only.

Cheremukhin, Anton, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinski (2017) "The Industrialization and Economic Development of Russia through the Lens of a Neoclassical Growth Model," *Review of Economic Studies*, Vol. 84, No. 2, pp. 613–649.

Cooper, Julian M. (2012) “Science-technology policy and innovation in the USSR,” slides, CEELBAS Workshop "Russia’s Skolkovo in Comparative and Historical Perspective", held on 12 June 2012 at UCL SSEES.

Criscuolo, Chiara, Ralf Martin, Henry Overman, and John Van Reenen (2012) “The causal effects of an industrial policy,” NBER Working Paper 17842, National Bureau of Economic Research.

Davis, Donald R. and David E. Weinstein (2002) “Bones, bombs, and break points: The geography of economic activity,” *American Economic Review*, Vol. 92, No. 5, pp. 1269–1289.

De Witt, Nicholas (1961) *Education and professional employment in the U.S.S.R.*, Washington, D.C.: National Science Foundation.

Dexter, Keith and Ivan Rodionov (2016) “The factories, research and design establishments of the Soviet defence industry: A guide. Version 17,” working paper, The University of Warwick, Department of Economics.

Diamond, Alexis and Jasjeet S. Sekhon (2013) “Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies,” *The Review of Economics and Statistics*, Vol. 95, No. 3, pp. 932–945.

Donaldson, Dave and Adam Storeygard (2016) “Thew view from above: Applications of satellite data in economics,” *Journal of Economic Perspectives*, Vol. 30, No. 4, pp. 171–198.

Drucker, Joshua M. (2012) “The spatial extent of agglomeration economies: Evidence from three U.S. manufacturing industries,”Technical Report CES-WP-12-01, US Census Bureau Center for Economic Studies.

Duranton, Gilles (2015) “Growing through Cities in Developing Countries,” *World Bank Research Observer*, Vol. 30, No. 1, pp. 39–73.

Duranton, Gilles and Diego Puga (2004) “Micro-foundations of urban agglomeration economies,” in Henderson, J. Vernon and Jacques-François Thisse eds. *Handbook of Regional and Urban Economics*, Vol. 4: Cities and Geography, Chap. 48, pp. 2063–2117.

von Ehrlich, Maximilian and Tobias Seidel (2018) “The persistent effects of place-based policy: Evidence from the West-German Zonenrandgebiet,” *American Economic Journal: Economic Policy*, Vol. 10, No. 4, pp. 344–374.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr (2010) “What causes industry agglomeration? Evidence from coagglomeration patterns,” *American Economic Review*, Vol. 100, No. 3, pp. 1195–1213.

Fan, Jingting and Ben Zou (2017) “Industrialization from scratch: The Third Front and local economic development in China’s hinterland,” Mimeo.

Friebel, Guido and Sergei Guriev (2005) “Attaching workers through in-kind payments: Theory and evidence from Russia,” *World Bank Economic Review*, Vol. 19, No. 2, pp. 175–202.

Ganguli, Ina (2015) “Immigration & ideas: What did Russian scientists ’bring’ to the US?” *The Journal of Labor Economics*, Vol. 33, No. S1, pp. S257–S288, Part 2, July.

Givord, Pauline, Rolande Rathelot, and Patrick Sillard (2013) “Place-based tax exemptions and displacement effects: An evaluation of the *Zones Franches Urbaines* program,” *Regional Science and Urban Economics*, Vol. 43, No. 1, pp. 151–163.

Glaeser, Edward L. and Joshua D. Gottlieb (2008) “The economics of place-making policies,” *Brookings Papers on Economic Activity*, pp. 155–239.

——— (2009) “The wealth of cities: Agglomeration economies and spatial equilibrium in the United States,” *Journal of Economic Literature*, Vol. 74, No. 4, pp. 983–1028.

Glaeser, Edward L., Hedi D. Kallal, José A. Scheinkman, and Andrei Schleifer (1992) “Growth in cities,” *Journal of Political Economy*, Vol. 100, No. 6, pp. 1126–1152.

Gokhberg, Leonid (1997) “Transformation of the Soviet R&D System,” in Gokhberg, L., M. J. Peck, and J. Gács eds. *Russian Applied Research and Development: Its Problems and its Promise*, Laxenburg, Austria: International Institute for Applied Systems Analysis, pp. 9–33.

Gorodnichenko, Yuriy, Jan Svejnar, and Katherine Terrell (2014) “When does FDI have positive spillovers? Evidence from 17 transition market economies,” *Journal of Comparative Economics*, Vol. 42, No. 4, pp. 954–969.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti (2010) "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings," *Journal of Political Economy*, Vol. 118, No. 3, pp. 536–598.

Gregory, Paul R. (2004) *The Political Economy of Stalinism: Evidence from the Soviet Secret Archives*: Cambridge University Press.

Gregory, Paul R. and Mark Harrison (2005) "Allocation under Dictatorship: Research in Stalin's Archives," *Journal of Economic Literature*, Vol. 43, No. 3, pp. 721–761.

Gregory, Paul R. and Andrei Markevich (2002) "Creating Soviet Industry: The House That Stalin Built," *Slavic Review*, Vol. 61, No. 4, pp. 787–814.

Gu, Xing Sam and Paul R. Rosenbaum (1993) "Comparison of multivariate matching methods: Structures, distances, and algorithms," *Journal of Computational and Graphical Statistics*, Vol. 2, No. 4, pp. 405–420.

Hall, Bronwyn H., Jacques Mairesse, and Pierre Mohnen (2010) "Measuring the returns to R&D," in Hall, B. H. and N. Rosenberg eds. *Handbook of the Economics of Innovation*: Elsevier.

Harrison, Mark (2018) "Secrecy and State Capacity: A Look Behind the Iron Curtain," mimeo.

Heblich, Stephan, Marlon Seror, Hao Xu, and Yanos Zylberbeg (2018) "Place-Based Policies in the Long Run: Evidence from Million-Rouble Plants in China," mimeo.

Ivanov, Denis S. (2016) "Human capital and knowledge-intensive industries location: Evidence from Soviet legacy in Russia," *Journal of Economic History*, Vol. 76, No. 3, pp. 736–768.

Jaffe, Adam B. (1989) "Real effects of academic research," *American Economic Review*, Vol. 79, No. 5, pp. 957–970.

Kline, Patrick M. and Enrico Moretti (2014a) "Local economic development, agglomeration economies, and the Big Push: 100 years of evidence from the Tennessee Valley Authority," *The Quarterly Journal of Economics*, Vol. 121, No. 1, pp. 275–331.

——— (2014b) "People, places, and public policy: Some simple welfare economics of local economic development programs," *Annual Review of Economics*, Vol. 6, pp. 629–662.

Kuboniwa, Masaaki (2011) "The Russian growth path and TFP changes in light of estimation of the production function using quarterly data," *Post-Communist Economies*, Vol. 23, No. 3, pp. 311–325.

Lappo, G. M. and P. M. Polyan (2008) "Naukograds of Russia: Yesterday's forbidden and semi-forbidden cities - today's growth points (Naukogrady Rossii: Vcherashniye zapretnyye i poluzapretnyye goroda - segodnyashniye tochki rosta)," *World of Russia (Mir Rossiyi)*, No. 1, pp. 20–49, Available in Russian only.

Lychagin, Sergey, Joris Pinkse, Margaret E. Slade, and John Van Reenen (2016) "Spillovers in space: Does geography matter?" *Journal of Industrial Economics*, Vol. 64, No. 2, pp. 295–335.

Marshall, Alfred (1890) *Principles of economics*, London: Macmillan, 8th edition.

Melo, Patricia C., Daniel J. Graham, and Robert B. Noland (2009) "A meta-analysis of estimates of urban agglomeration economies," *Regional Science and Urban Economics*, Vol. 39, No. 3, pp. 332–342.

Mikhailova, Tatiana (2012a) "Gulag, WWII and the Long-run Patterns of Soviet City Growth," in MPRA Paper 41758: University Library of Munich, Germany.

——— (2012b) "Where Russians should live: A counterfactual alternative to Soviet location policy," in MPRA Paper 35938: University Library of Munich, Germany.

Moretti, Enrico (2004) "Workers' education, spillovers and productivity: Evidence from plant-level production functions," *American Economic Review*, Vol. 94, No. 3, pp. 656–690.

——— (2011) "Local labor markets," in *Handbook of Labor Economics*: Elsevier.

——— (2013) "Real wage inequality," *American Economic Journal: Applied Economics*, Vol. 5, No. 1, pp. 65–103.

Moretti, Enrico, Claudia Steinwender, and John Van Reenen (2016) "The intellectual spoils of war? Defense R&D, productivity and spillovers," mimeo.

NAS (2002) *Successes and difficulties of small innovative firms in Russian nuclear cities: Proceedings of a Russian-American workshop. Committee on Small Innovative Firms in Russian Nuclear Cities, Office for Central Europe and Eurasia Development, Security, and Cooperation, National Research Council, in cooperation with the Institute of Physics and Power Engineering, Obninsk, Russia*, Washington, D.C.: National Academy of Sciences.

Neumark, David and Jed Kolko (2010) "Do enterprise zones create jobs? Evidence from California's Enterprise Zone Program," *Journal of Urban Economics*, Vol. 68, No. 1, pp. 1–19.

Neumark, David and Helen Simpson (2014) "Place-based policies," NBER Working Paper 20049, National Bureau of Economic Research.

OECD and Statistical Office of the European Communities (2005) *Oslo Manual: Guidelines for collecting and interpreting innovation data, 3rd edition, The measurement of scientific and technological activities*, Luxembourg: OECD Publishing, pp.162.

Roback, Jennifer (1982) "Wages, rents and the quality of life," *Journal of Political Economy*, Vol. 90, No. 6, pp. 1257–1278.

Rosen, Sherwin (1979) "Wage-based indexes of urban quality of life," in Miezkowski, P. N. and M. R. Straszheim eds. *Current Issues in Urban Economics*: Johns Hopkins University Press, Baltimore, MD.

Rosenbaum, Paul R. (2002) *Design of observational studies*: Springer.

Rowland, Richard H. (1996) "Russia's Secret Cities," *Post-Soviet Geography and Economics*, Vol. 37, No. 7, pp. 426–462.

Ruchnov, K. M. and E. G. Zaitseva (2011) "Crisis of Russian science cities (Krizis naukogradov Rossiyi),"Technical report, mimeo, Available in Russian only.

Saltykov, Boris G. (1997) "The reform of Russian science," *Nature*, Vol. 388, No. 6637, pp. 16–18.

Siddiqi, Asif A. (2008) "Soviet Space Power During the Cold War," in Gillespie, Paul G. and Grant T. Weller eds. *Harnessing the Heavens: National Defense Through Space*: Colorado Springs: United States Air Force Academy, pp. 134–150.

——— (2015) "Scientists and Specialists in the Gulag: Life and Death in Stalin's Sharashka.," *Kritika: Explorations in Russian and Eurasian History*, Vol. 16, No. 3, pp. 557–588.

Sinani, Evis and Klaus E. Meyer (2004) "Spillovers of technology transfer from FDI: the case of Estonia," *Journal of Comparative Economics*, Vol. 32, No. 3, pp. 445–466.

State Bank (1946) *List of branches of the State Bank of the USSR as of 1st March 1946 (Spisok kontor, otdeleniy i agentstv Gosudarstvennogo banka SSSR po sostoyaniyu na 1-ye marta 1946 goda)*, Moscow: State Bank of the USSR, Available in Russian only.

Syverson, Chad (2011) "What determines productivity?" *Journal of Economic Literature*, Vol. 49, No. 2, pp. 326–365.

Vernadsky State Geological Museum and United States Geological Survey, 20010600 (2001) "rails.shp. Railroads of the Former Soviet Union," U.S. Geological Survey, Denver, CO.

Zhao, Zhong (2004) "Using matching to estimate treatment effects: Data requirements, matching metrics, and Monte Carlo evidence," *The Review of Economics and Statistics*, Vol. 86, No. 1, pp. 91–107.

Appendices

A Russian R&D and scientists over the transition

Figures A.1 and A.2 illustrate the magnitude of the shock that the dissolution of the USSR has represented for Russian science, and the later developments. Our data sources are: Gokhberg (1997), the Russian Statistical Yearbooks for various years, and the OECD Main Science and Technology Indicators (MSTI) database.

Figure A.1: Gov. spending in R&D as a share of GDP, Russia vs. OECD, 1989-2010

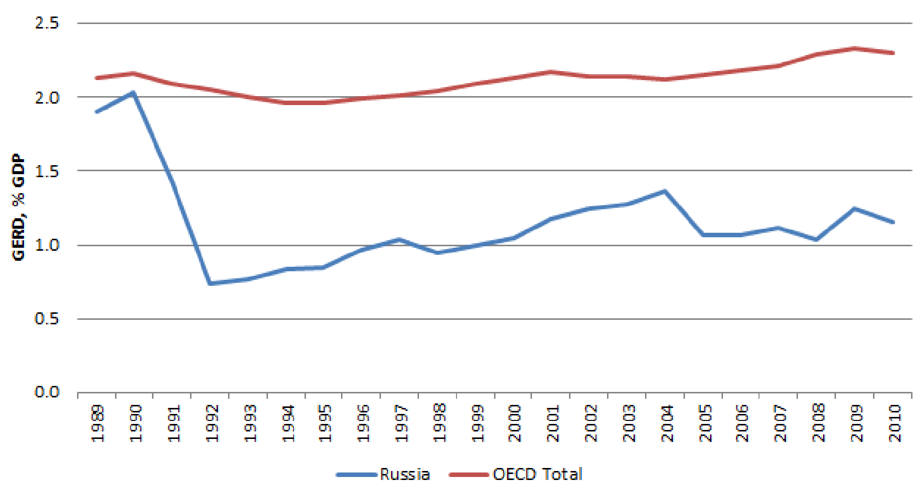
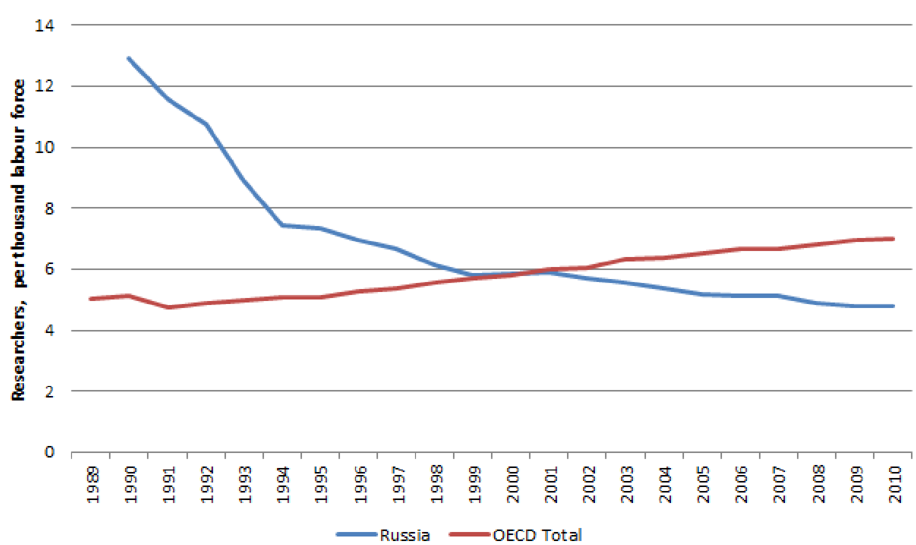


Figure A.2: Share of scientists in the labor force, Russia vs. OECD, 1989-2010



B Science Cities in the Soviet Union and Russia

Table B.1: Science Cities

				Naukograd			Closed city[d]		Priority specialization areas[a]							
No.	Location[a]	Oblast	Founded[e]	Year Soviet status[e]	Year Russian status[f]	Type[a]	Past	Now	Military base	Air rocket and space research	Electronics and radio engineering	Automation, IT and instrumentation	Chemistry, chemical physics and new materials	Nuclear complex	Energetics	Biology, biotechnology and agricultural sciences
1	Biysk	Altai Krai	1718	1957	2005	1	No	No	No	No	No	No	Yes	No	No	Yes
2	Mirny	Arkhangelsk	1957	1966		2	Yes	Yes	No	Yes	No	No	No	No	No	No
3	Severodvinsk	Arkhangelsk	1936	1939		2	Yes	No	Yes	No	No	Yes	No	No	No	No
4	Znamensk	Astrakhan	1948	1962		2	Yes	Yes	Yes	Yes	No	No	No	No	No	No
5	Miass	Chelyabinsk	1773	1955		1	No	No	No	Yes	No	Yes	No	No	No	No
6	Ozyorsk	Chelyabinsk	1945	1945		2	Yes	Yes	No	No	No	No	No	Yes	No	No
7	Snezhinsk	Chelyabinsk	1957	1957		2	Yes	Yes	No	No	No	No	No	Yes	No	No
8	Tryokhgorny	Chelyabinsk	1952	1952		2	Yes	Yes	No	No	No	Yes	No	Yes	No	No
9	Ust-Katav	Chelyabinsk	1758	1942		1	No	No	No	Yes	No	Yes	No	No	No	No
10	Kaspiysk[c]	Dagestan	1932	1936		2	No	No	No	No	No	Yes	No	No	No	No
11	Akademgorodok	Irkutsk	1949	1988		5	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12	Angarsk[c]	Irkutsk	1948	1957		2	No	No	No	No	No	No	Yes	Yes	No	No
13	Obninsk	Kaluga	1946	1946	2000	2	No	No	No	No	No	Yes	No	Yes	Yes	No
14	Sosensky[c]	Kaluga	1952	1973		3	No	No	No	No	No	Yes	No	No	No	No
15	Komsomolsk-on-Amur	Khabarovsk	1932	1934		1	No	No	No	Yes	No	Yes	No	No	No	No
16	Krasnodar-59[b]	Krasnodar					No	No		No	No	No	No	No	No	No

[a] Based on Agirrechu (2009), unless specified otherwise. Type: 1 - Science Cities, "scientific core" established in existing cities, which often had a particular historical significance; 2 - Science Cities that received city status simultaneously (or a few years later) with the creation of a scientific or scientific-industrial complex at a practically new location (often in the "open field"); 3 - Science Cities that have arisen in existing settlements and received city status after obtaining scientific functions; 4 - Science Cities that do not have city status; 5 - academic town.

[b] Based on NAS (2002).

[c] Based on Lappo and Polyan (2008).

[d] Russian Wikipedia article on closed cities (ZATO), 28 September 2016.

[e] Wikipedia articles for each city, 28 September 2016.

[f] Russian Wikipedia article on Science Cities, 28 September 2016.

Table B.1 – continued from previous page

				Naukograd			Closed city[d]			Priority specialization areas[a]							
No.	Location[a]	Oblast	Founded[e]	Year Soviet status[e]	Year Russian status[f]	Type[a]	Past	Now	Military base	Air rocket and space research	Electronics and radio engineering	Automation, IT and instrumentation	Chemistry, chemical physics and new materials	Nuclear complex	Energetics	Biology, biotechnology and agricultural sciences
17	Akademgorodok	Krasnoyarsk	1944	1965		5	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
18	Zelenogorsk	Krasnoyarsk	1956	1956		2	Yes	Yes	No	No	No	No	Yes	Yes	No	No
19	Zheleznogorsk	Krasnoyarsk	1950	1954		2	Yes	Yes	No	Yes	No	No	No	Yes	No	No
20	Kurchatov[c]	Kursk	1968	1976		2	No	No	No	No	No	No	No	Yes	Yes	No
21	Gatchina	Leningrad	1928	1956		1	No	No	No	No	No	Yes	No	Yes	No	No
22	Primorsk	Leningrad	1268	1948		1	No	No	No	Yes	No	No	No	No	No	No
23	Sosnovy Bor	Leningrad	1958	1962		3	Yes	Yes	No	No	Yes	No	No	Yes	Yes	No
24	Zelenograd	Moscow City	1958	1958		2	No	No	No	No	Yes	Yes	No	No	No	No
25	Avtopoligon	Moscow Oblast	1964	1964		4	No	No	No	No	No	Yes	No	No	No	No
26	Balashikha	Moscow Oblast	1830	1942		1	No	No	No	Yes	No	Yes	No	No	No	No
27	Beloozersky	Moscow Oblast	1961	1961		4	No	No	No	Yes	No	No	No	No	No	No
28	Chernogolovka	Moscow Oblast	1710	1956	2008	3	No	No	No	No	No	Yes	Yes	No	No	No
29	Dolgoprudny	Moscow Oblast	1931	1951		2	No	No	No	No	No	Yes	Yes	No	No	No
30	Dubna	Moscow Oblast	1956	1956	2001	2	No	No	No	Yes	No	Yes	No	Yes	No	No
31	Dzerzhinsky	Moscow Oblast	1938	1956		3	No	No	No	Yes	No	Yes	Yes	No	No	No
32	Fryazino	Moscow Oblast	1584	1953	2003	3	No	No	No	Yes	Yes	Yes	No	No	No	No
33	Istra	Moscow Oblast	1589	1946		1	No	No	No	Yes	No	Yes	No	No	Yes	No
34	Khimki	Moscow Oblast	1850	1950		1	No	No	No	Yes	No	Yes	No	No	No	No
35	Klimovsk	Moscow Oblast	1882	1940		1	No	No	No	No	No	Yes	No	No	No	No

[a] Based on Agirrechu (2009), unless specified otherwise. Type: 1 - Science Cities, "scientific core" established in existing cities, which often had a particular historical significance; 2 - Science Cities that received city status simultaneously (or a few years later) with the creation of a scientific or scientific-industrial complex at a practically new location (often in the "open field"); 3 - Science Cities that have arisen in existing settlements and received city status after obtaining scientific functions; 4 - Science Cities that do not have city status; 5 - academic town.

[b] Based on NAS (2002).

[c] Based on Lappo and Polyan (2008).

[d] Russian Wikipedia article on closed cities (ZATO), 28 September 2016.

[e] Wikipedia articles for each city, 28 September 2016.

[f] Russian Wikipedia article on Science Cities, 28 September 2016.

Table B.1 – continued from previous page

No.	Location[a]	Oblast	Founded[e]	Naukograd: Year Soviet status[e]	Naukograd: Year Russian status[f]	Naukograd: Type[a]	Closed city[d]: Past	Closed city[d]: Now	Priority specialization areas[a]: Military base	Air rocket and space research	Electronics and radio engineering	Automation, IT and instrumentation	Chemistry, chemical physics and new materials	Nuclear complex	Energetics	Biology, biotechnology and agricultural sciences
36	Korolyov	Moscow Oblast	1938	1946	2001	1	No	No	No	Yes	No	Yes	Yes	No	No	No
37	Krasnoarmeysk	Moscow Oblast	1928	1934		1	No	No	No	No	No	Yes	Yes	No	No	No
38	Krasnogorsk[c]	Moscow Oblast	1932	1942		3	No	No	No	No	Yes	No	No	No	No	No
39	Krasnoznamensk	Moscow Oblast	1950	1950		2	Yes	Yes	No	Yes	No	No	No	No	No	No
40	Lukhovitsy	Moscow Oblast	1594	1957?		1	No	No	No	Yes	No	No	No	No	No	No
41	Lytkarino	Moscow Oblast	1939	1957		1	No	No	No	Yes	No	Yes	No	No	No	No
42	Lyubertsi[c]	Moscow Oblast	1623	1948		1	No	No	No	No	No	Yes	No	No	No	No
43	Mendeleyevo	Moscow Oblast	1957	1965		4	No	No	No	No	No	Yes	No	No	No	No
44	Mytishchi[c]	Moscow Oblast	1460	1935		1	No	No	No	No	No	Yes	No	No	No	No
45	Obolensk	Moscow Oblast	1975	1975		4	Yes	No	No	No	No	No	No	No	No	Yes
46	Orevo	Moscow Oblast	1954	1954		4	No	No	No	Yes	No	No	No	No	No	No
47	Peresvet	Moscow Oblast	1948	1948		2	No	No	No	Yes	No	No	No	No	No	No
48	Protvino	Moscow Oblast	1960	1960	2008	3	Yes	No	No	No	No	Yes	No	Yes	No	No
49	Pushchino	Moscow Oblast	1956	1966	2005	3	No	No	No	No	No	No	No	No	No	Yes
50	Remmash	Moscow Oblast	1957	1957		4	No	No	No	Yes	No	No	No	No	No	No
51	Reutov	Moscow Oblast	1492-1495	1940	2003	1	No	No	No	Yes	No	Yes	No	No	No	No
52	Tomilino	Moscow Oblast	1894	1961		4	No	No	No	Yes	No	Yes	No	No	No	No
53	Troitsk	Moscow Oblast	1617	1977	2007	3	No	No	No	No	No	Yes	No	Yes	No	No
54	Yubileyny	Moscow Oblast	1939	1950		3	Yes	No	No	Yes	No	No	No	No	No	No

[a] Based on Agirrechu (2009), unless specified otherwise. Type: 1 - Science Cities, “scientific core” established in existing cities, which often had a particular historical significance; 2 - Science Cities that received city status simultaneously (or a few years later) with the creation of a scientific or scientific-industrial complex at a practically new location (often in the “open field”); 3 - Science Cities that have arisen in existing settlements and received city status after obtaining scientific functions; 4 - Science Cities that do not have city status; 5 - academic town.

[b] Based on NAS (2002).

[c] Based on Lappo and Polyan (2008).

[d] Russian Wikipedia article on closed cities (ZATO), 28 September 2016.

[e] Wikipedia articles for each city, 28 September 2016.

[f] Russian Wikipedia article on Science Cities, 28 September 2016.

Table B.1 – continued from previous page

No.	Location[a]	Oblast	Founded[e]	Naukograd: Year Soviet status[e]	Naukograd: Year Russian status[f]	Type[a]	Closed city[d]: Past	Closed city[d]: Now	Priority specialization areas[a]: Military base	Air rocket and space research	Electronics and radio engineering	Automation, IT and instrumentation	Chemistry, chemical physics and new materials	Nuclear complex	Energetics	Biology, biotechnology and agricultural sciences
55	Zheleznodorozhny	Moscow Oblast	1861	1952		3	No	No	No	No	No	Yes	No	No	No	No
56	Zhukovsky	Moscow Oblast	1933	1947	2007	2	No	No	No	Yes	No	No	No	No	No	No
57	Zvyozdny gorodok	Moscow Oblast	1960	1960		4	Yes	Yes	No	Yes	No	No	No	No	No	No
58	Apatity (Akademgorodok)	Murmansk	1926	1954		2	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
59	Polyarnye Zori[c]	Murmansk	1968	1973		2	No	No	No	No	No	No	No	Yes	Yes	No
60	Balakhna (Pravdinsk)	Nizhny Novgorod	1932	1941		1	No	No	No	No	Yes	No	No	No	No	No
61	Dzerzhinsk	Nizhny Novgorod	1606	1930		2	No	No	No	No	No	No	Yes	No	No	No
62	Sarov	Nizhny Novgorod	1310	1947		1	Yes	Yes	No	No	No	No	No	Yes	No	No
63	Akademgorodok	Novosibirsk	1957	1957		5	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
64	Koltsovo	Novosibirsk	1979	1979	2003	4	Yes	No	No	No	No	No	No	No	No	Yes
65	Krasnoobsk	Novosibirsk	1970	1978		4	No	No	No	No	No	No	No	No	No	Yes
66	Novosibirsk-49[b]	Novosibirsk					No	No		No	No	No	No	No	No	No
67	Omsk-5[b]	Omsk					No	No								
68	Zarechny	Penza	1954	1958		2	Yes	Yes	No	No	No	No	No	Yes	Yes	No
69	Perm-6[b]	Perm					No	No								
70	Bolshoy Kamen[c]	Primorsk Krai	1947	1954		2	Yes	Yes	No	No	No	Yes	No	No	No	No
71	Volgodonsk[c]	Rostov	1950	1976		3	No	No	No	No	No	No	No	Yes	Yes	No
72	Zernograd	Rostov	1929	1935		1	No	No	Yes	No	No	No	No	No	No	Yes
73	Petergof	Saint Petersburg	1711	1960	2005	1	No	No	No	No	No	Yes	Yes	No	No	No

[a] Based on Agirrechu (2009), unless specified otherwise. Type: 1 - Science Cities, "scientific core" established in existing cities, which often had a particular historical significance; 2 - Science Cities that received city status simultaneously (or a few years later) with the creation of a scientific or scientific-industrial complex at a practically new location (often in the "open field"); 3 - Science Cities that have arisen in existing settlements and received city status after obtaining scientific functions; 4 - Science Cities that do not have city status; 5 - academic town.

[b] Based on NAS (2002).

[c] Based on Lappo and Polyan (2008).

[d] Russian Wikipedia article on closed cities (ZATO), 28 September 2016.

[e] Wikipedia articles for each city, 28 September 2016.

[f] Russian Wikipedia article on Science Cities, 28 September 2016.

Table B.1 – continued from previous page

				Naukograd			Closed city[d]		Priority specialization areas[a]							
No.	Location[a]	Oblast	Founded[e]	Year Soviet status[e]	Year Russian status[f]	Type[a]	Past	Now	Military base	Air rocket and space research	Electronics and radio engineering	Automation, IT and instrumentation	Chemistry, chemical physics and new materials	Nuclear complex	Energetics	Biology, biotechnology and agricultural sciences
74	Desnogorsk[c]	Smolensk	1974	1982		2	No	No	No	No	No	No	No	Yes	Yes	No
75	Lesnoy	Sverdlovsk	1947	1954		2	Yes	Yes	No	No	No	No	No	Yes	No	No
76	Nizhnyaya Salda	Sverdlovsk	1760	1958		1	No	No	No	Yes	No	No	No	No	No	No
77	Novouralsk	Sverdlovsk	1941	1949		2	Yes	Yes	No	No	No	No	Yes	Yes	No	No
78	Verhnyaya Salda[c]	Sverdlovsk	1778	1933		3	No	No	No	No	No	No	Yes	No	No	No
79	Zarechny	Sverdlovsk	1955	1955		2	No	No	No	No	No	No	No	Yes	Yes	No
80	Michurinsk	Tambov	1635	1932	2003	1	No	No	No	No	No	Yes	No	No	No	Yes
81	Zelenodolsk[c]	Tatarstan	1865	1949		1	No	No	No	No	No	Yes	No	No	No	No
82	Akademgorodok	Tomsk	1972	1972		5	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
83	Seversk	Tomsk	1949	1949		2	Yes	Yes	No	No	No	No	Yes	Yes	No	No
84	Redkino	Tver	1843	1950		4	No	No	No	No	Yes	No	Yes	No	No	No
85	Solnechny	Tver	1947	1951		4	Yes	Yes	Yes	No	No	Yes	No	No	No	No
86	Udomlya[c]	Tver	1478	1984		3	No	No	No	No	No	No	No	Yes	Yes	No
87	Glazov[c]	Udmurtia	1678	1948		1	No	No	No	No	No	No	No	Yes	No	No
88	Votkinsk[c]	Udmurtia	1759	1957		1	No	No	No	Yes	No	Yes	No	No	No	No
89	Dimitrovgrad	Ulyanovsk	1698	1956		1	No	No	No	No	No	No	No	Yes	Yes	No
90	Kovrov	Vladimir	1778	1916		1	No	No	No	No	No	Yes	No	No	No	No
91	Melenki	Vladimir	1778			1	No	No	No	No	No	Yes	No	No	No	No
92	Raduzhny	Vladimir	1971	1971		2	Yes	Yes	No	No	Yes	Yes	No	No	No	No

[a] Based on Agirrechu (2009), unless specified otherwise. Type: 1 - Science Cities, "scientific core" established in existing cities, which often had a particular historical significance; 2 - Science Cities that received city status simultaneously (or a few years later) with the creation of a scientific or scientific-industrial complex at a practically new location (often in the "open field"); 3 - Science Cities that have arisen in existing settlements and received city status after obtaining scientific functions; 4 - Science Cities that do not have city status; 5 - academic town.

[b] Based on NAS (2002).

[c] Based on Lappo and Polyan (2008).

[d] Russian Wikipedia article on closed cities (ZATO), 28 September 2016.

[e] Wikipedia articles for each city, 28 September 2016.

[f] Russian Wikipedia article on Science Cities, 28 September 2016.

Table B.1 – continued from previous page

				Naukograd			Closed city[d]		Priority specialization areas[a]							
No.	**Location[a]**	**Oblast**	**Founded[e]**	**Year Soviet status[e]**	**Year Russian status[f]**	**Type[a]**	**Past**	**Now**	**Military base**	**Air rocket and space research**	**Electronics and radio engineering**	**Automation, IT and instrumentation**	**Chemistry, chemical physics and new materials**	**Nuclear complex**	**Energetics**	**Biology, biotechnology and agricultural sciences**
93	Novovoronezh[c]	Voronezh	1957	1964		2	No	No	No	No	No	No	No	Yes	Yes	No
94	Borok	Yaroslavl	1807	1956		4	No	No	No	No	No	No	No	No	No	Yes
95	Pereslavl-Zalessky	Yaroslavl	1152	1964		1	No	No	No	No	No	Yes	Yes	No	No	No

[a] Based on Agirrechu (2009), unless specified otherwise. Type: 1 - Science Cities, “scientific core” established in existing cities, which often had a particular historical significance; 2 - Science Cities that received city status simultaneously (or a few years later) with the creation of a scientific or scientific-industrial complex at a practically new location (often in the “open field”); 3 - Science Cities that have arisen in existing settlements and received city status after obtaining scientific functions; 4 - Science Cities that do not have city status; 5 - academic town.

[b] Based on NAS (2002).

[c] Based on Lappo and Polyan (2008).

[d] Russian Wikipedia article on closed cities (ZATO), 28 September 2016.

[e] Wikipedia articles for each city, 28 September 2016.

[f] Russian Wikipedia article on Science Cities, 28 September 2016.

C Municipal level data sources and variables

Table C.1: Municipal level data sources and variables

Data type	Data sub-type	Data source	Description
		Factors guiding the selection of location of Science Cities	
Administrative	Various identification information for municipality, region and federal district	OpenStreetMaps, available through GIS-LAB (`http://gis-lab.info/qa/osm-adm.html`)	Unique municipality, federal district and region (oblast, krai, republic) identificators, codes and names
Population	1959 census data	January 1959 Soviet Census, available through Demoscope (`http://demoscope.ru/weekly/ssp/census.php?cy=3`)	All population in municipality in 1959, estimates for some municipalities
Geography	Area	Calculated in QGIS based on OpenStreetMaps	Municipality area calculated in QGIS, measured in square kilometers
	Coordinates of the municipality center		GPS coordinates of the center of municipality calculated in QGIS
	Altitude	CGIAR.[a] Consortium for Spatial Information (CGIAR-CSI) SRTM 90m Digital Elevation Data, version 4, available at `http://srtm.csi.cgiar.org/`	Altitude of municipality in meters (mean, median, SD, min and max value)
	Temperatures in January and July	WorldClim version 1 (`http://www.worldclim.org/version1`), developed by Hijmans et al. (2005)	Monthly temperature data, for the period 1960-90, assigned to municipalities in QGIS. Average, median, standard deviation, minimum, and maximum.

Continued on next page

Table C.1 – Continued from previous page

Data type	Data sub-type	Data source	Description
	Railroad	Vernadsky State Geological Museum and United States Geological Survey, 20010600 (2001) and Central management unit of the military communications of the Red Army (1943)	Data on railroads were constructed using railroads shapefile describing the railroads of the former Soviet Union as of the early 1990s prepared by Vernadsky State Geological Museum and United States Geological Survey, 20010600 (2001), along with a map of railroads from 1943 from Central management unit of the military communications of the Red Army (1943) to manually remove any differences between the situation depicted in the shapefile and the 1943 map. Indicator equal to 1 if municipality has access to railroad, and 0 otherwise. Railroads are as of late 1940s.
	Coastline/major river/lake	Natural Earth, 1:10m Physical Vectors version (`http://www.naturalearthdata.com/downloads/10m-physical-vectors/`)	Indicator equal to 1 if municipality has access to coast/major river/lake and 0 otherwise.
	Distances	Calculated in QGIS based on the sources specified above	Distance (in km) from the center of municipality to the nearest railroad, coast, major river, lake, USSR border, plant (of any type), and HEI (of any type).
Level of industrial development	Data on the factories, research and design establishments of the Soviet defence industry in 1947	Dexter and Rodionov (2016). The dataset contains almost 30,000 entries and includes the name, location, main branch of defence production, establishment type as well as the start and end date for the establishment's military work.	Number of factories (*zavod*) and subordinated organizations.
R&D institutes			Number of Scientific Research & Design Institutes (*NII*, *TsNII*, and *GSPI*), design bureaus, and test sites.
Graduate share institutes (HEI)	HEIs in the municipality in 1959	De Witt (1961)	Number of all HEIs, HEIs specialisng in technical sciences, and HEIs specialising in biology and medical sciences

Continued on next page

Table C.1 – Continued from previous page

Data type	Data sub-type	Data source	Description
State Bank	State bank State Bank branches in 1946	Bircan and De Haas (2020), originally in State Bank (1946)	Number of State Bank branches
		Long-term outcomes of interest	
Patents	Applications to EPO	European Patent Office. Patents are matched to municipalities via inventors' addresses.	Number of patents applications to EPO in 2006-15, by inventor (simple and fractional counting)
Population	2010 census data	2010 Russian census, available at `http://www.gks.ru/free_doc/new_site/perepis2010/croc/perepis_itogi1612.htm`	All population, population with higher education, and population with PhD or doctoral degrees in municipality in 2010
Night time lights	Average stable night lights	Version 4 DMSP-OLS Nighttime Lights Time Series, National Oceanic and Atomspheric Administration (NOAA) (`http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html`)	Average night lights for 1992-94 and 2009-11, cleaned of gas flares
SMEs	Results of the 2010 SME census	Rosstat (Federal State Statistics Service) (`http://www.gks.ru/free_doc/new_site/business/prom/small_business/itog-spn.html`)	The dataset contains information on the number of firms, revenue, number of employees, fixed assets, and fixed capital investment by size, 1- and 4-digit ISIC sector. We use SMEs per capita, SME sales per worker (labor productivity), and SME sales per unit of fixed labor, all sectors and manufacturing only. The SME census does not cover ZATOs, so 16 Science Cities, which are also ZATOs, are not covered.
Municipal budget	Average municipal budget revenues and expenditures over 2006-16	Rosstat (Federal State Statistics Service)	Average annual municipal budget revenues and expenditures over 2006-16, in 2010 prices. Total values and breakdown by major categories.

[a] CGIAR is a global partnership of research organizations dedicated to reducing poverty and hunger, improving human health and nutrition, and enhancing ecosystem resilience through agricultural research. CGIAR-CSI is spatial science community that facilitates CGIAR's international agricultural development research using spatial analysis, GIS, and remote sensing: `http://www.cgiar-csi.org/`.

D Testing the dynamics of the night lights measures

This appendix has two related purposes. The first is to provide a regression-based formal test to the claim advanced in section 3.3.7, that the trends of night lights of treated and control observations from the matched sample do not converge over time. To clear this concern, Table D.1 below reports estimates of the following simple regression model:

$$Y_{it} = \pi_0 T_t + \pi_1 S_i \cdot T_t + \tau_t + \varepsilon_{it}, \tag{D.1}$$

where Y_{it} is a night lights measure, T_t is a linear trend, S_i is a Science City dummy, τ_t is a year effect and ε_{it} is an error term which is allowed to be autocorrelated in time. Clearly, parameter π_1 represents differences in the linear trend between Science Cities and their matched counterparts. The estimates in Table D.1 are obtained from two different samples: the one comprising all matched Science Cities alongside their paired controls (column 1) and the subsample restricted to the "historical" Science Cities and their matches (column 2). Standard errors are calculated via the Newey-West HAC formula, allowing autocorrelation of ε_{it} up to 10 years (virtually the same results are obtained with longer time windows). The estimates of π_1, albeit small in magnitude, are positive and statistically significant; this suggests that, if anything, Science Cities are on a divergent trend of economic development (as proxied by the night lights measure).

Table D.1: Estimates of parameters π_0 and π_1 from model (D.1)

	All Science Cities (1)	Historical Science Cities (2)
Linear trend, all (π_0)	0.0074***	0.0052***
	(0.0001)	(0.0001)
Trend difference for S.C.s (π_1)	0.0002**	0.0001
	(0.0001)	(0.0001)
Year effects	YES	YES
Number of observations	3,002	2,470

Notes: *, ** and *** denote significance at the 10, 5, and 1 per cent level, respectively. Standard errors reported in parentheses are estimated with the Newey-West HAC formula, allowing for autocorrelation up to 10 years.

Having ensured that there is no convergence, one can move to the estimation of a "grand ATT" on the night lights panel data, which is the second purpose of this appendix. This is useful, since the separate ATT estimation the ATT on a single year often delivers results that are not statistically significant, which is possibly due to statistical noise; a "grand ATT" estimation that appropriately exploits the longitudinal dimension of the night lights data can circumvent this problem. This task is undertaken by estimating the following simple model:

$$Y_{it} = \psi S_i + \tau'_t + \varepsilon'_{it}. \tag{D.2}$$

Here, parameter ψ represents the causal effect of Science City status, and identification is ensured by restricting the analysis to the matched sample. It is reasonable to expect the estimate(s) of ψ to be positive, but not necessarily statistically significant when allowing for autocorrelated disturbances. Observe that this exercise cannot be performed along with the estimation of separate linear trends for Science Cities and their matches, because it would result in data overfitting. Once again, the autocorrelation of the error term is allowed for up to 10 years. The results are reported in Table D.2.

Table D.2: Estimates of parameter ψ from model (D.2)

	All Science Cities		Historical Science Cities	
	(1)	(2)	(3)	(4)
Science City (ψ)	0.3040*	0.3521**	0.2944	0.3992**
	(0.1639)	(0.1751)	(0.1844)	(0.1967)
Year effects	YES	YES	YES	YES
Bias-adjusted estimate	NO	YES	NO	YES
Number of observations	3,154	3,154	2,622	2,622

Notes: *, ** and *** denote significance at the 10, 5, and 1 per cent level, respectively. Standard errors reported in parentheses are estimated with the Newey-West HAC formula, allowing for autocorrelation up to 10 years.

The baseline estimate reported in column 1 amounts to about one third of the standard deviation of the night lights measure, which is similar to the ATT estimate of the 2009-11 average from Tables 4-5, and is statistically significant at the 10 per cent level. The estimate in column 2 is instead obtained by adding bias adjustment terms as per Abadie and Imbens (2006, 2011) to the left-hand side measures Y_{it}, allowing the linear correction model to vary by year. This results in an even higher measure of ψ – up to half the measure's standard deviation – which is significant at the 5 per cent level. By restricting the analysis to the "historical" Science Cities, one obtains similar if less precise estimates; they are reported in columns 3 and 4 (without and with bias adjustment, respectively). To summarize, by looking at night lights as a proxy of economic activity, it appears that the effect of Science Cities is statistically significant and constant over time: yet another result that offers little support to the mean-reversion hypothesis.

E Effects by categories Science City: visual representation

This appendix provides a more extensive, visual representation of the treated-control differences used to perform the t-tests discussed in section 3.3.8 and reported in Table 9. Specifically, for each categorization and each key outcome of interest, a bar chart is reported; in these figures, each bar represents a matched pair, and their (possibly negative) heights represent the pair's treated-control difference. Furthermore, the bars are arranged in ascending order and colored by one of the two split categories. Hence, an uneven distribution of colors along the horizontal axis is indicative of differences between categories, and vice versa.

Figure E.1: Differences in total population, "secret" vs. "usable" Science Cities

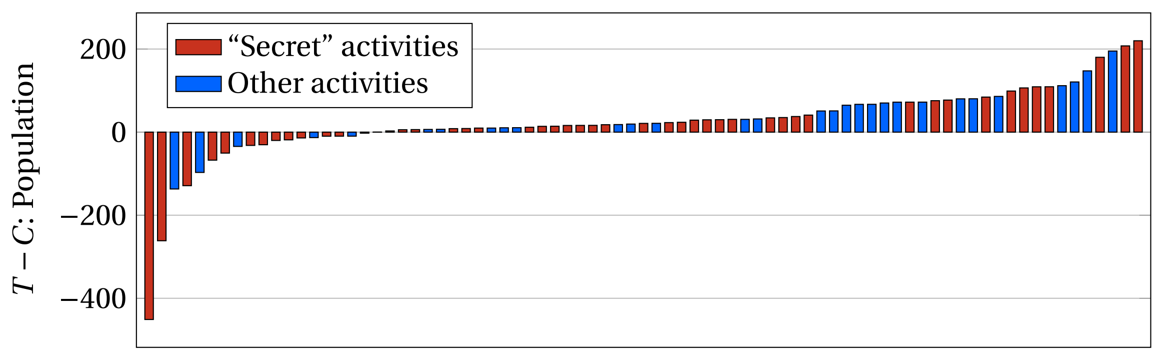


Figure E.2: Differences in the graduate share, "secret" vs. "usable" Science Cities

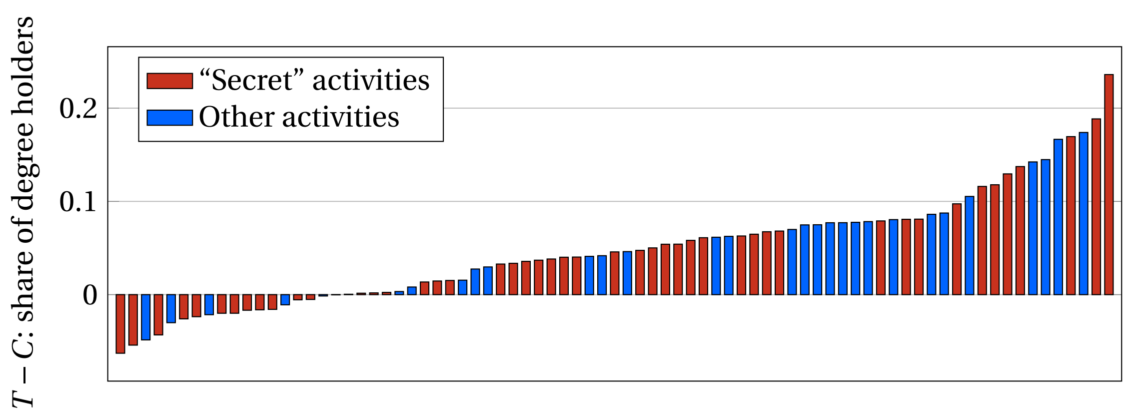


Figure E.3: Differences in total patent output, “secret” vs. “usable” Science Cities

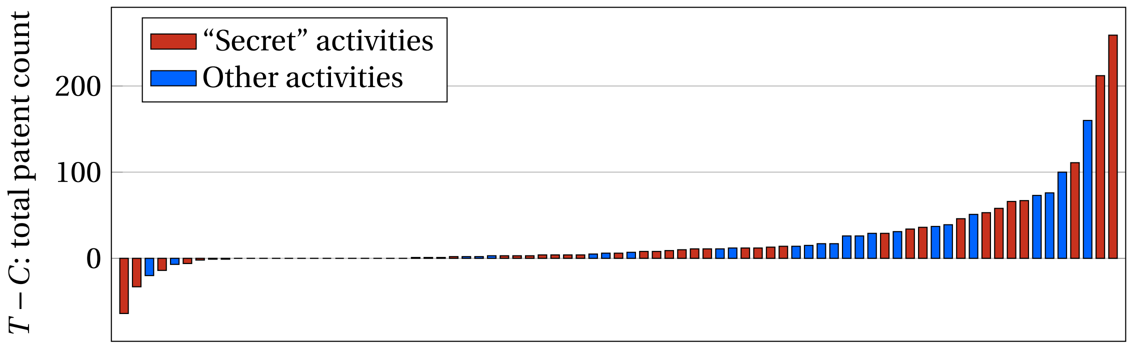


Figure E.4: Differences in night lights (2009-11), “secret” vs. “usable” Science Cities

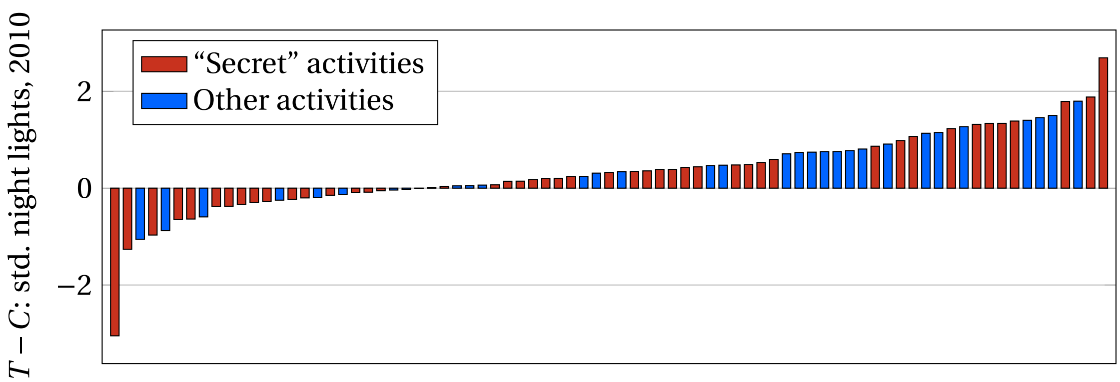


Figure E.5: Differences in total population, Science Cities built from scratch vs. pre-existing

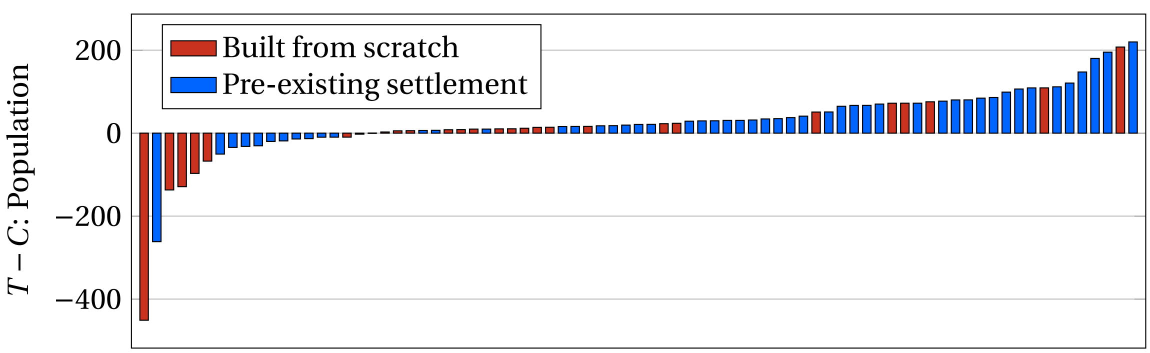


Figure E.6: Differences in the graduate share, Science Cities built from scratch vs. pre-existing

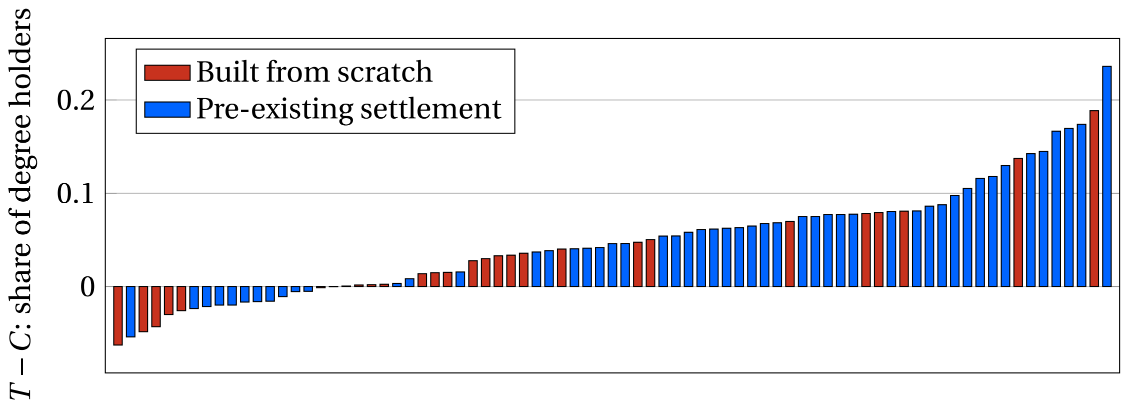


Figure E.7: Differences in total patent output, Science Cities built from scratch vs. pre-existing

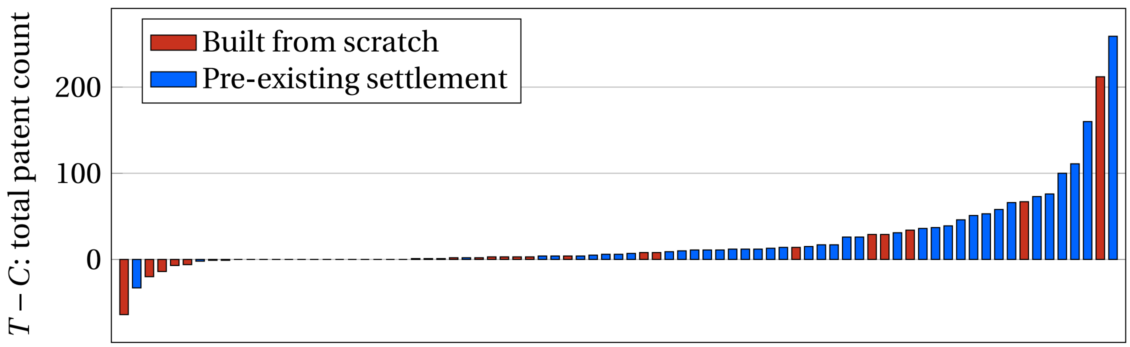


Figure E.8: Differences in night lights (2009-11), Science Cities built from scratch vs. pre-existing

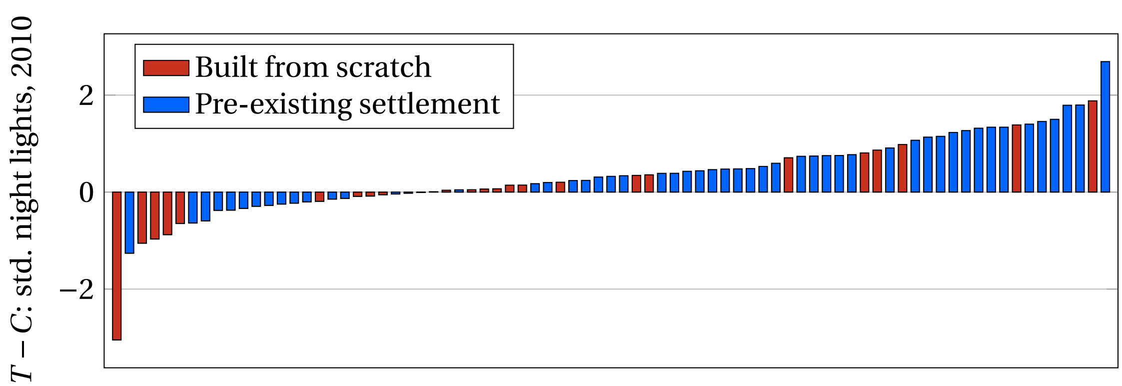


Figure E.9: Differences in total population, Science Cities from Agirrechu's list vs. others

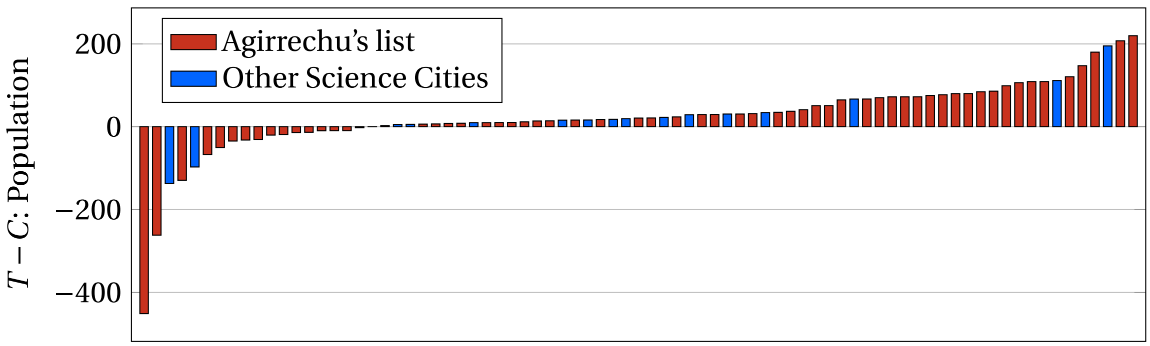


Figure E.10: Differences in the graduate share, Science Cities from Agirrechu's list vs. others

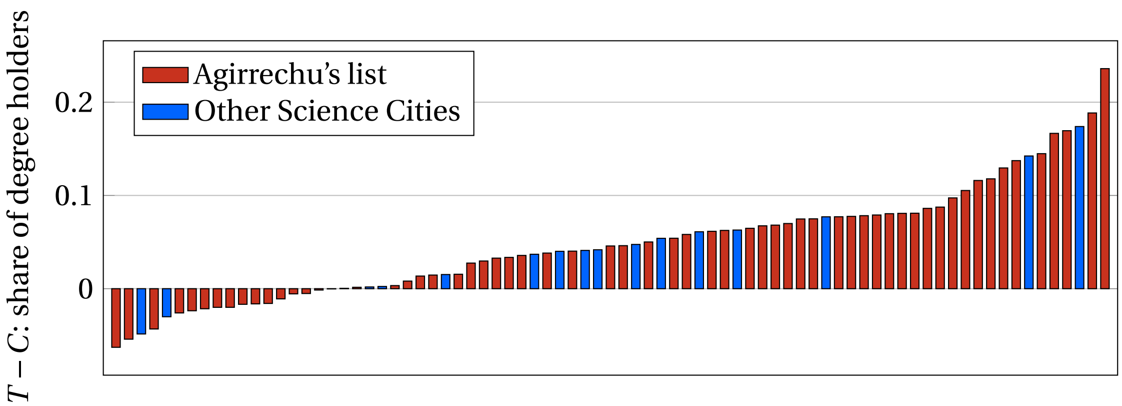


Figure E.11: Differences in total patent output, Science Cities from Agirrechu's list vs. others

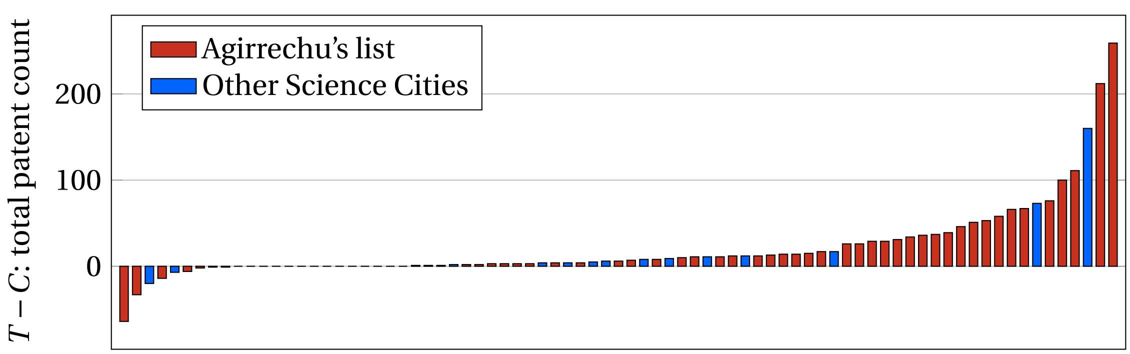
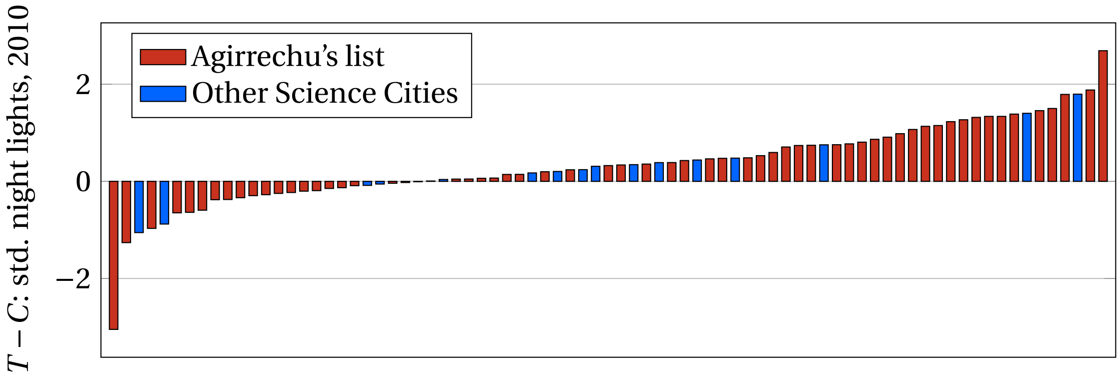


Figure E.12: Differences in night lights (2009-11), Science Cities from Agirrechu's list vs. others



F Aggomeration variables descriptive statistics

Table F.1: Agglomeration potential measures ($\lambda = 3$): Descriptives

	Mean	Std. dev.
Fractional patents, SC	0.0273	0.4557
Graduate share (%), SC	0.0529	0.6788
Postgratuate share (%), SC	0.0015	0.0220
R&D employment share (%), SC	0.0105	0.1367
Fractional patents, non-SC	0.2877	5.6412
Graduate share (%), non-SC	0.1565	0.8735
Postgratuate share (%), non-SC	0.0043	0.0251
R&D employment share (%), non-SC	0.0150	0.0933

Table F.2: Agglomeration potential measures ($\lambda = 3$): Correlations

	Fractional patents, SC	Graduate share (%), SC	Postgraduate share (%), SC	R&D employment share (%), SC	Fractional patents, non-SC	Graduate share (%), non-SC	Postgraduate share (%), non-SC
Graduate share (%), SC	0.6911***						
Postgratuate share (%), SC	0.5652***	0.9430***					
R&D employment share (%), SC	0.8340***	0.9459***	0.8436***				
Fractional patents, non-SC	-0.0031	-0.004	-0.0034	-0.0039			
Graduate share (%), non-SC	-0.0107	-0.0139	-0.0118	-0.0138	0.1584***		
Postgratuate share (%), non-SC	-0.0103	-0.0134	-0.0113	-0.0132	0.1492***	0.9816***	
R&D employment share (%), non-SC	-0.0096	-0.0125	-0.0106	-0.0123	0.1829***	0.9550***	0.9684***

G BEEPS V Russia

BEEPS is an enterprise survey, the objective of which is to gain an understanding of firms' perceptions of the environment in which they operate in order to be able to assess the constraints to private sector growth and enterprise performance. It covers topics related to infrastructure, sales and supplies, degree of competition, land and permits, crime, finance, business-government relations, labor and establishment performance. BEEPS is implemented by private contractors, using face-to-face interviews in the country's official language(s). In BEEPS V, for the first time 37 Russian regions were covered, at least one in each federal district. The survey was primarily targeted at top managers (CEOs), but in reality the respondents often included accountants or operations managers. A total of 4,220 face-to-face interviews were completed, on average 114 interviews per region (see Table G.1).

Table G.1: BEEPS V Russia sample breakdown

Region	Number of interviews	Region	Number of interviews
Central	***1124***	***Siberian***	***709***
Belgorod	120	Irkutsk	131
Kaluga	121	Kemerovo	124
Kursk	87	Krasnoyarsk	89
Lipetsk	121	Novosibirsk	123
Moscow City	121	Omsk	120
Moscow Oblast	122	Tomsk	122
Smolensk	71	***Southern***	***328***
Tver	120	Krasnodar	88
Voronezh	121	Rostov	120
Yaroslavl	120	Volgograd	120
Far Eastern	***334***	***Urals***	***199***
Khabarovsk	122	Chelyabinsk	79
Primorsky Krai	120	Sverdlovsk	120
Sakha (Yakutia)	92	***Volga***	***922***
North Caucasian	***120***	Bashkortostan	106
Stavropol Krai	120	Kirov	134
Northwestern	***484***	Mordovia	120
Kaliningrad	122	Nizhni Novgorod	82
Leningrad	120	Perm	120
Murmansk	120	Samara	120
St. Petersburg	122	Tatarstan	120
		Ulyanovsk	120
Total			***4,220***

Sources: EBRD-World Bank BEEPS V Russia.

Also for the first time, BEEPS V Russia included an innovation module, with the aim of obtaining a better understanding of innovation–not only product innovation, but also process, organization and marketing innovation, as well as R&D and protection of innovation. The main questionnaire contained questions that determined eligibility for participation in the innovation module, which was based on the third edition of the Oslo Manual OECD and Statistical Office of the European Communities (2005). The so-called filtering questions were asked with the help of show cards, which contained examples of the relevant innovations to facilitate a common understanding of the definition of innovation. While non-innovators did not receive additional questions on innovations, innovating firms were asked to provide more information, including a detailed description of their main product or process innovation (in terms of impact on sales or costs respectively). Firms were only asked the relevant parts of the innovation module, which in turn collected more detailed information on how the firms innovate, the level of innovativeness and how important innovation is for the firms, as well as on R&D spending and patents. Firms were asked to specify their main innovative product and process. In Russia, 86.3 per cent of Innovation Module interviews were completed face-to-face immediately after the main questionnaire; 9.9 per cent were completed during a follow-up phone call, and the rest during a second face-to-face visit or immediately after completing section H in the main questionnaire.

The detailed descriptions of the firms' main product or process innovation were used to analyze whether the respective innovation complied with the formal definitions of product and process innovation, taking into account the firm's main business. Based on this assessment, innovators could be reclassified as non-innovators, or moved to another category of innovation than the one self-reported. As a result, only 51.9 per cent of Russian companies that said they introduced new products did product innovation and only 59.7 per cent of Russian companies that said they introduced new processes met the definition of process innovation. The cleaning of innovations can only be done for product or process, that is, technological innovations, as no additional questions were asked for non-technological innovations. We also corrected the indicator for R&D spending in the last three years based on the answers in the innovation module. There was a significant variation across regions on all of these measures, which could reflect both the competence of interviewers as well as understanding of the respondents.