European Bank
for Reconstruction and Development

The long-run effects of R&D place-based policies: evidence from Russian science cities

Helena Schweiger, Alexander Stepanov and Paolo Zacchia

Abstract
We study the long-run effects of historical place-based policies targeting research and development (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 political and military or strategic considerations. 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. We find that in the modern Russian economy, 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 that have to do with the differential use of public resources, and we find limited support for a case of equilibrium reversion. Lastly, by analysing firm-level data we obtain evidence in favour of spillover effects with a wide spatial breadth.

Keywords: Place-based policy, R&D, science cities.

JEL Classification: O30, R11, N94.

Contact details: Helena Schweiger, European Bank for Reconstruction and Development, One Exchange Square, London, EC2A 2JN, UK. Email: `schweigh@ebrd.com`.

Helena Schweiger and Alexander Stepanov are at the EBRD; Paolo Zacchia is at the IMT Lucca School of Advanced Studies.

We would like to thank Charles Becker, Ralph de Haas, Sergei Guriev, Maria Gorban, Denis Ivanov, Sergei Izmalkov, Patrick Kline, Olga Kuzmina, Andrei Markeevich, Enrico Moretti, Gérard Roland and Natalya Volchkova for helpful discussions, as well as the participants at the AEA 2018, 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 and seminars at the EBRD, Higher School of Economics, New Economic School, U.C. Berkeley, and University of Genoa 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.

Working Paper No. 216 **Prepared in June 2018**

1 Introduction

The effectiveness of public support to 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 usually motivated by the existence of 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 characterised 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 localised innovation policies mixes with the one about broader (that is, not innovation-specific) *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]

This paper examines the long-run impact of a localised innovation policy: the establishment of highly specialised science cities in the territory of modern Russia during Soviet times. These are 95 middle-sized urban centres that were created or developed by the Soviet government according to a strategic plan of technological advancement. Science cities hosted a high concentration of R&D facilities – often the only driving economic activity in town – typically built around a specific technological purpose. Since science cities emerged in the context of technological and military competition in the Cold War, most of them were, unsurprisingly, specialised 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.

Our contribution to the existing literature is four-fold. First, we collect a rich municipal-level dataset, covering geographic, historical and present characteristics of Russian municipalities. Second, we estimate the impact of science cities at the level of municipalities and firms. Third, we uncover the channels that create persistence and disentangle the role of government transfers from agglomeration economies. Fourth, we examine the demographic and economic dynamics to make statements about the persistence of differences in socio-economic outcomes of interest.

While one may question whether the institutional context of Russian science cities is

[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).

comparable to that of other industrialised countries, this historical experience stands out with some unique features that motivate its analysis. First, the locations of science cities were typically chosen by the Soviet leadership with criteria that are unusual for a capitalistic market economy. According to Aguirrechu (2009), since the Soviet government had the power to allocate both physical and human capital where it deemed necessary, the potential for economic development and local human capital accumulation was typically not, *at the margin*, a determinant of a location's choice for the establishment of a science city. Instead, the choice between any two places that were similarly suited to host such a settlement usually fell on the one that offered better secrecy and safety from foreign interference (in the form of R&D espionage), or that satisfied other political and strategic criteria. This greatly diminishes concerns for selection biases due to unobserved determinants of future development, which typically affect studies about innovative clusters in other countries.

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; only recently was it partially resumed for 14 of the former towns, which today bear the official name of *Naukogrady* (Russian for "science cities"). Together, these historical developments indicate that both the selection into the science cities programme and the timing of the policy's discontinuation were largely driven by exogenous factors, orthogonal to determinants of current demographic and economic conditions. In addition, by analysing 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.

We estimate the effect of the past establishment of a science city on the following set of present characteristics of Russian municipalities: human capital (measured as the share of the population with either graduate or postgraduate qualifications), innovation (evaluated in terms of patent output) 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. Our main identifying assumption is that, conditional on these 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 both historical and more recently observed local characteristics.

Our results can be summarised as follows. In today's Russia, 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 programme'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 60 per cent. We also perform a more in-depth analysis of demographic outcomes and night lights proxy for economic development, which reveals no evidence of mean reversion.

We interpret our 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 provides different predictions about several city-level outcomes to the extent that science cities are inherently better places to live, workers' mobility is more or less restricted, the initial allocation modified individual preferences for location, or agglomeration forces such as knowledge spillovers exist. In light of these predictions, we interpret our empirical results about the productivity and wages of highly skilled workers, which can be explained in equilibrium either by systematic productivity differentials or by agglomeration forces, as suggestive of localised knowledge spillovers at work. Under the hypotheses of our empirical strategy, idiosyncratic productivity shocks are statistically unlikely to explain the observed differences.

This interpretation contrasts with the study by von Ehrlich and Seidel (2015, vES) 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 analysing municipal budget data. We find that, with equal available resources, at least some former science cities spend more per capita on physical infrastructure (such as roads) with respect to matched towns. In our context, we see the mechanism proposed by vES as possibly complementary to the agglomeration forces. Arguably, science cities emerged from the planned economy with a better endowment of both human and physical capital. We conjecture that a more skilled population was instrumental in maintaining such advantages over and beyond the transition.

Furthermore, we complement the municipal-level empirical analysis with an additional set of estimates based on firm-level data. We use data about 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. BEEPS V is particularly useful in this regard, as it features an innovation module with detailed information about recent innovative activities by firms. This analysis is meant to evaluate whether in the modern Russian economy, the effect of science cities spills over onto other firms that are located nearby, and to what economic and geographical extent. We see these results as reinforcing our conclusion that the municipal-level differentials are at least in part caused by knowledge spillovers, since firms are observed to be more likely to conduct R&D when they are based relatively close to science cities.

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). The majority of these papers analyse 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, 2015). Among the few that focus, like us, on a non-western, formerly planned economy, the notable contributions are Wang (2013) on Chinese Special Economic Zones and Fan and Zou (2017) on the "Third Front" state-driven industrialisation of inner China. The empirical challenges faced by these studies are typically about constructing appropriate control groups, and disentangling the policy's direct effects from spillovers. Our paper is most directly related to the study by Kline and Moretti (2014a) on the Tennessee Valley Authority and that by Fan and Zou (2017) on China's Third Front. While these contributions uncover long-run effects from historical place-based policies focusing on physical capital, the science cities programme stands out as it was mostly about an investment in knowledge. The similarities extend to the empirical strategy, as we also exploit unique historical circumstances of political 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, localised 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.[2] Recently, the issue has caught the attention of economists working in more diverse fields. Moretti (2004) shows that in US cities, the level of education of the workforce affects firm productivity across sectors. Ellison et al. (2010) simultaneously test all three Marshallian theories by looking at the co-location of plants across industries. Greenstone et al. (2010) demonstrate the existence of local productivity spillovers following the opening of a "Million Dollar Plant". In two separate contributions, Bloom et al. (2013) and Lychagin et al. (2016) find an association between firms' R&D spending and the productivity of those nearby.

The specific institutional setting of this paper relates it to other, somewhat diverse, contributions on the consequences of historically massive forms of government intervention in 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 (2012) evaluates negative welfare effects from the regional demographic policies enacted by the Soviet Union. However, the picture looks 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 in the OECD countries, increases in government-funded R&D for military purposes have positive net effects on total factor productivity (TFP), despite crowding out private expenditures in R&D.

This paper is organised as follows. Section 2 briefly introduces the history and characteristics of Soviet science cities. Section 3 outlines the conceptual framework of the paper. Section 4

[2] 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.

focuses on the long-run effects of R&D at the municipal level: it outlines the empirical methodology, describes the data employed and discusses the empirical results. Section 5 examines the long-run effects of R&D at the firm level and is similarly structured. Section 6 concludes.

2 Historical and institutional background

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).[3] Science cities emerged in this environment. They were 95 middle-sized urban centres which the Soviet government endowed with a high concentration of research and development facilities, each devoted to a particular scientific and technical specialisation.[4] Science cities began to develop around strategically important (military) research centres from the mid-1930s;[5] however, the majority of them were established after the Second World War, especially in the 1950s.

As they specialised 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 centres, 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. As we detail later in section 4.1, the choice of science cities' locations was typically guided by military, political and other idiosyncratic considerations (Aguirrechu, 2009). For the sake of providing better incentives to individuals who worked in science cities, the Soviet government 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 (Aguirrechu, 2009).

Starting in the 1940s, with the need to protect the secrecy of the nuclear weapons programme in the Cold War environment (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 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

[3] The Soviet innovation system is briefly described in Appendix A.

[4] The term "science city" (*Naukograd*) was first introduced in 1991 (Ruchnov and Zaitseva, 2011). The former Soviet Union was not a 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 labour camp system. 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.

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,[6] a fact even more dramatic as Russian GDP shrank by about 50 per cent in the initial years of the transition. As a consequence of much lower wages, total employment in R&D also fell by about 50 per cent.[7] 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 programmes 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. Our analysis of recent municipal budgets (see section 4.3) confirms that science cities receive today, if anything, *lower* governmental transfers than comparable towns, especially if modern *Naukogrady* are removed from the sample.

[6]We calculated these figures using as sources: Gokhberg (1997), the Russian Statistical Yearbooks for various years, and the OECD Main Science and Technology Indicators (MSTI) database.

[7]Whereas in Soviet times the wages of scientists were 10-20 per cent higher than the average, they dropped to 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.

3 Analytical framework

To facilitate the interpretation of the long-run effect of the science cities programme, we use a spatial equilibrium framework standard in the urban economics literature. Specifically, we adapt the model by Moretti (2011, 2013) which itself extends Rosen (1979), Roback (1982) and Glaeser and Gottlieb (2008, 2009).

3.1 Model set-up

Consider two ex-ante identical cities, s and z, which could be inhabited by different types of workers: those of high educational level or "skill", and those of relatively lower skill. This dichotomous classification is typically interpreted in terms of differences in higher educational achievement. In this context, highly skilled workers can be even more narrowly identified 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 logarithm of the mass of highly skilled workers employed in city $c \in s, z$ at time t as h_{ct}, while ℓ_{ct} is the corresponding notation for low-skilled workers.

At time $t = 0$, the two cities are part of the Soviet Union which, for exogenous reasons, attributes the status of science city only to s. We interpret the science cities programme 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 s, so that $(h_{s0} - h_{z0}) > 0$. At the same time, we assume that $(\ell_{s0} - \ell_{z0}) \leq 0$, which reflects the spatial segregation of economic activity in the USSR. We also consider the possibility that the urban planning choices and the public investments enacted in science cities might have made them more enjoyable locations to live in. In urban economics parlance one would say that the amenities a_s of science city s are higher than the amenities a_z of the ordinary locality z: hence $\tilde{a} \equiv a_s - a_z \geq 0$.

At time $t = 1$, the two cities are part of modern Russia, a market economy, and workers of both types self-select into either location. Following Moretti (2011, 2013) 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, amenities and idiosyncratic preferences:[8]

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

where w_{nc} is the log-wage earned by workers of type n in city c, while e_{nic} denotes the

[8]Typically, in these models workers' utility also depends – negatively – on city-specific price indexes r_c. For simplicity, here 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 labour supply and agglomeration effects.

idiosyncratic taste of individual i for city c. The individuals' relative preferences for the two localities are distributed as follows:

$$e_{nis} - e_{niz} \sim \mathcal{U}\left[-m_n + b_n, m_n + b_n\right]. \tag{2}$$

For both types $n = h, \ell$, m_n represents the overall degree of mobility of workers of type n, while b_n is the type-specific average bias towards science city s. Intuitively, the higher m_n, the lower the importance of idiosyncratic tastes for the choice of location.

In Moretti (2011, 2013), it is maintained that $b_h = b_\ell = 0$. However, here we assume that the two groups' preferences are asymmetric:

$$\begin{aligned} b_h &= b\left(h_{s0} - h_{z0}\right) > 0 \\ b_l &= b\left(\ell_{s0} - \ell_{z0}\right) \leq 0, \end{aligned} \tag{3}$$

where $b(\cdot)$ is an increasing monotone function with $b(0) = 0$. This hypothesis introduces a mechanism of path-persistence: if an individual 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 $t = 0$.[9] Another interpretation of (3) is in terms of restrictions on mobility: in Russia, internal mobility used to be very costly, if not altogether impossible, due to regulation inherited from Soviet times.[10] This can be represented as a group differential in the average moving cost.

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 this was largely a simplification meant to abstract from the degree of substitutability between skills in Moretti's analysis, this characteristic of the model can be given a contextual interpretation here: if workers of type h are researchers, type h firms correspond with 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:

$$\begin{aligned} y_{hc} &= x_{hc} + \theta_h h_c + \mu h_c + \left(1-\mu\right) k_{hc} \\ y_{\ell c} &= x_{\ell c} + \theta_\ell h_c + \mu \ell_c + \left(1-\mu\right) k_{\ell c}, \end{aligned} \tag{4}$$

where x_{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 labour 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.

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

[9] Allowing $b_h, b_\ell \neq 0$ is omothetic to letting the value of amenities vary by worker type, as in Moretti. In this institutional context, it is important to make a mechanism of path-persistence in location choice explicit in our conceptual framework.

[10] A system of internal visas was in place until the early 2000s. Studies about internal migration rates in Russia in the 1990s show that they were very low (Andrienko and Guriev, 2004; Friebel and Guriev, 2005).

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 $\theta_\ell > 0$. Such a distinction between "restricted" and "general" spillover effects is, to the best of our knowledge, new in theoretical frameworks of urban economics. The model provides different equilibrium predictions to the extent that $\theta_h > 0$, $\theta_\ell > 0$, or both – with corresponding empirical implications.

3.2 Spatial equilibrium

In a spatial equilibrium at $t = 1$, a marginal worker of either type must be indifferent between cities s and z. This implies that the supply of, say, highly skilled labour in either city is determined by the following condition (we drop timing subscripts for convenience):

$$m_h \left(\frac{h_s - h_z}{\overline{h}} \right) = w_{hs} - w_{hz} + \tilde{a} + b_h, \tag{5}$$

where $\overline{h} \equiv h_s + h_z$ is given and, if $\theta_h > 0$, is also such that $\overline{h} < \theta_h^{-1} \mu m_h$.[11] The equilibrium wage differentials $(w_{hs} - w_{hz})$ are obtained as the difference between the marginal productivity of highly skilled labour in the two cities; this difference, in turn, depends on the equilibrium in the capital market.[12] A symmetric analysis applies to the case of low-skilled labour.

As a result, the relative difference in equilibrium highly skilled employment between the two cities can be expressed as:

$$(h_s - h_z) = \frac{\left[\tilde{x}_h + \mu(\tilde{a} + b_h)\right] \overline{h}}{\mu m_h - \theta_h \overline{h}} \geq 0, \tag{6}$$

where $\tilde{x}_h \equiv x_{hs} - x_{hz}$ is the difference in log-TFP of type h firms between the two cities. Equation (6) is interpreted as follows: there are three forces that cause science cities to continue hosting a larger number of researchers and highly skilled workers after the transition. These are: (i) inherent productivity differentials ($\tilde{x}_h > 0$), (ii) superior amenities in science cities ($\tilde{a} > 0$), and (iii) path-dependence mechanisms ($b_h > 0$). All these forces are stronger the more highly skilled workers are mobile (lower m_h) and the larger are the agglomeration effects (larger θ_h). Importantly, agglomeration effects alone are not sufficient to cause employment differentials, at least in the equilibrium under analysis: they only complement factors (i-iii) that affect the supply of labour.

[11] This condition is necessary to avoid that the denominators of (6) and (7) turn negative, breaking their interpretability. In practice, spillovers θ_h and the total mass of log-researchers $\overline{h}$ cannot be simultaneously "too high," or the equilibrium would degenerate into full spatial concentration of highly skilled workers.

[12] Equilibrium in the capital market implies that the marginal productivity of capital must be equal in the two cities: $(k_{hs} - k_{hz}) = (h_s - h_z) + \mu^{-1} \tilde{x}_h$. The difference between the inverse labour demands in the two cities can be expressed as: $(w_{hs} - w_{hz}) = \mu^{-1} \left[\tilde{x}_h + \theta_h (h_s - h_z)\right]$.

The relative difference in the productivity of highly skilled workers equals that of their wages:

$$\left(y_{hs} - y_{hz}\right) - (h_s - h_z) = (w_{hs} - w_{hz}) = \frac{m_h \tilde{x}_h + \theta_h \overline{h}\,(\tilde{a} + b_h)}{\mu m_h - \theta_h \overline{h}}. \tag{7}$$

This result bears some important implications for our empirical analysis. First, absent agglomeration forces ($\theta_h = 0$), these differences are proportional to the log-TFP differentials $\tilde{x}_h$. Second, if the latter are null ($\tilde{x}_h = 0$), any positive difference in the productivity and wages of highly skilled workers between science cities and comparable locations is indicative of increasing returns.[13] In the empirical analysis, we measure the difference in municipal-level outcomes, observed about 20 years following the dissolution of the USSR, between several dozens of science cities and their matched counterparts. Thus, by standard statistical arguments it is unlikely that exogenous shocks to TFP alone could explain any systematic productivity or wage differentials for highly skilled workers.

For low-skilled workers, the equilibrium log-employment difference reads (for given $\overline{\ell} \equiv \ell_s + \ell_z$) as:

$$(\ell_s - \ell_z) = \frac{\overline{\ell}}{m_\ell}\left[\frac{\tilde{x}_\ell + \theta_\ell\,(h_s - h_z)}{\mu} + \tilde{a} + b_\ell\right] \gtreqless 0, \tag{8}$$

and its sign is undetermined. In fact, path-persistence mechanisms that *may* push low-skilled workers away from science cities ($b_\ell \leq 0$) could be more than compensated by: amenity differentials ($\tilde{a} \geq 0$), TFP differentials ($\tilde{x}_\ell \equiv x_{\ell s} - x_{\ell z} \geq 0$) and, if science cities host more highly skilled workers, cross-sector agglomeration forces ($\theta_\ell\,(h_s - h_z) \geq 0$). The equilibrium differentials in productivity and wages for low-skilled workers are:

$$\left(y_{\ell s} - y_{\ell z}\right) - (\ell_s - \ell_z) = (w_{\ell s} - w_{\ell z}) = \frac{\tilde{x}_\ell + \theta_\ell\,(h_s - h_z)}{\mu}. \tag{9}$$

Hence, by a reasoning analogous to the one outlined in the case of highly skilled workers, any empirical difference in those variables – in sectors unrelated to R&D – is evidence favourable to the operation of "generalised" spillover effects ($\theta_\ell > 0$).[14]

[13] Intuitively, under constant returns to scale ($\theta_h = 0$), the endogenous response of capital would equalise differences across the two cities in both the marginal and the average product of (highly skilled) labour, even in the presence of employment differentials.

[14] All these results would still hold, 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 labour mobility in equilibrium. See Moretti (2011, 2013) for an analysis of this model featuring negative externalities but without positive agglomeration forces.

4 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.

4.1 Empirical methodology

We compare the long-run outcomes Y_{iq} of municipalities hosting science cities with those of other, ordinary, municipalities which were similar in terms of geographical and socio-economic characteristics X_{ik} in the years following the Second World War, when the majority of science cities were established. $i = 1, \dots, N$ indexes municipalities; $q = 1, \dots, Q$ our long-run outcomes of interest; and $k = 1, \dots, K$ 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.

Our identifying assumption is that, conditional on the observed geographical and historical characteristics, the establishment of science cities did not depend on factors that would affect future outcomes. The Conditional Independence Assumption is motivated by the unique history of science cities: according to historical research on the topic, the selection criteria for the location of science cities were typically informed by idiosyncratic functional and military-strategic criteria that had often little to do with strict economic considerations. In particular, Aguirrechu (2009, p. 21) distinguishes between two groups of science cities according to their function and type of geographical location.[15] 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 *akademgorodki* were organised (e.g. in Novosibirsk, Tomsk, etc.)". Cities in the first group "hosted mainly organisations focusing on theoretical research". Instead – he continues – "[c]ities of the second group were located in the most remote areas of the country (although in densely populated regions), far away from large urban centres, 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."

In the case of science cities in the second group identified by Aguirrechu, it is easier to argue that military and intelligence factors were the main drivers of location choice. In this respect, the establishment of those science cities bears similarities to China's "Third Front" industrialisation policy examined by Fan and Zou (2017), which was also impelled by strategic considerations. At the extreme, considerations of this kind overrode all the others. In particular, science cities specialising in applied R&D fields such as the production of nuclear and strategic arms faced a much higher threat of bombing and spying, and were located in regions far from the borders and in municipalities far from the regional centre (with limited transport links) and

[15] Given the nature of the period during which most science cities were established and the associated political context, we were unable to access any systematic, reliable official information on how their locations were chosen.

previously poorly populated. Examples include Sarov and Snezhinsk.[16]

A parallel argument can also be extended to science cities of the first group: while less military, their driving location criteria were not less idiosyncratic. The majority of science cities of the first group – and about one third of the total – are located close to Moscow, because of their vicinity to "the Academy of Science, the All-Union Academy of Agricultural Sciences, the Academy of Medical Sciences, and some institutes subordinate to ministries" (Aguirrechu, 2009, p. 21). The location of the remaining science cities in the first group depended instead on historical and political circumstances of a different sort. Following the evacuation of factories from the European part of the Soviet Union beyond the Urals during the Second World War, those areas developed rapidly. This explains the concentration of many science cities close to industrial centres in the Urals. In addition, this propelled the establishment of a particular class of science cities, the so-called *akademgorodki* ("academic towns"), in Siberian centres to the east of the Urals with rising industrial and strategic importance but limited scientific capacities (Aguirrechu, 2009).[17]

That fact that the choice of location for science cities was largely driven by political and military factors does not mean that it was *unconditionally* orthogonal to economic outcomes. In fact, the quoted passages by Aguirrechu underline the fact that science cities of both groups were located in densely populated parts of Russia. Figure 1 illustrates this fact by showing the location of science cities superimposed on the chloropleth map of Russian regions distinguished by their population density. With some exceptions, science cities were established in the areas of Russia that were the most industrialised, urbanised, and with a more educated population, so that they could have easier access to qualified personnel and specialised inputs or were able to attract them at minor costs.[18]

Other geographical factors that might have affected both the location of science cities and future economic outcomes depended on the specialisation of science cities: heavy industry and nuclear technology need large amounts of water for their operations, therefore science cities focused on those areas were typically built close to rivers or lakes; analogously, science cities devoted to military shipbuilding had to be located on the coast. Based on the previous discussion, we argue

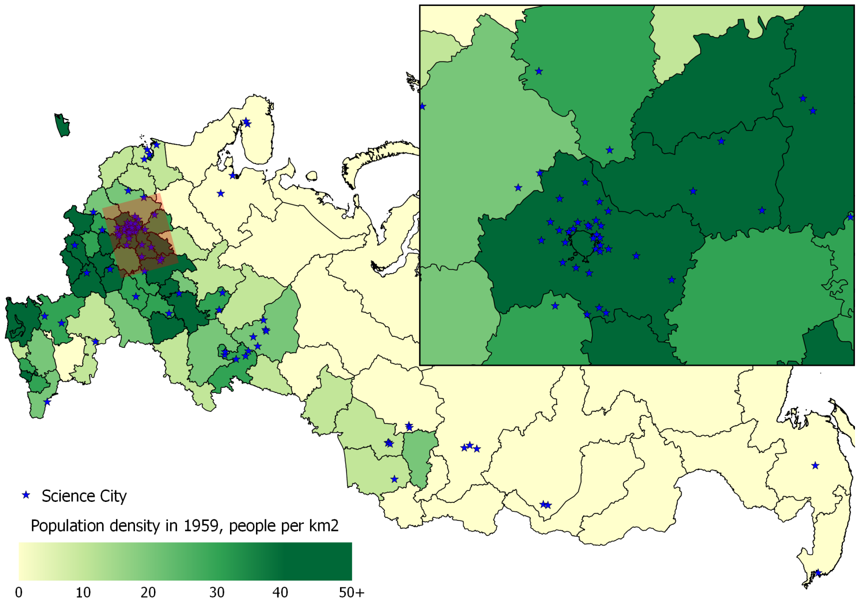
[16]These two places provide a particularly indicative example of idiosyncratic factors affecting the location of science cities: sometimes, this was determined by the presence of other science cities, or lack thereof. Snezhinsk (Chelyabinsk region) was established as a double of Sarov (Nizhny Novgorod 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 pattern of interplay between decisions affecting different science cities was not unique; for example, the four places specialised in production of enriched uranium were also located far from each other.

[17]Academic towns were semi-isolated neighbourhoods 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.

[18]For this reason, science cities are also for the most part located in the western, warmer part of Russia, within the humid continental climatic region typified by large seasonal temperature differences. 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; thus, for two localities with the same latitude, the eastern one is typically colder.

that *conditional* on the levels of economic development at the time of selection and some key geographical characteristics, science city status is independent of current socio-economic outcomes, because it was driven by military and political factors *at the margin*. Our strategy of matching on *observed* historical and geographical covariates is predicated on this idea.

Figure 1: Location of science cities and regional population density



Source: Table B.1 and ROSSTAT.

Our matching algorithm of choice is Mahalanobis matching, by which a science city s is matched to the ordinary municipality z with the lowest *Mahalanobis Distance* m_{sz}:

$$m_{sz}(\mathrm{x}_{is},\mathrm{x}_{iz}) = (\mathrm{x}_{is}-\mathrm{x}_{iz})^{\mathrm{T}}\,\Sigma\,(\mathrm{x}_{is}-\mathrm{x}_{iz}), \tag{10}$$

where x_{ic} is the vector of the K observable covariates for municipality i of type $c \in s, z$; while Σ is the empirical covariance matrix of the covariates. Matching is performed with replacement, so that a control municipality can be linked to multiple treated cities; in addition, we impose exact matching on two dummy variables: access to inland water and closed city status. Importantly, we match science cities that were subject to the "closed city" status described in section 2, to other non-science cities with similar restrictions (typically, these are places hosting military bases but lacking R&D content). By including geographical coordinates and variables that control for the local density of population and economic activity into vector x_{ic}, we make sure that science cities are matched to control towns that are remarkably close in space, especially in the densely populated parts of Russia such as the Moscow region. In so doing, we relieve concerns about the presence of spatially correlated unobservables affecting our results.

We obtain a unique association of treated-control observations which is based on the original set of 88 science cities in our dataset (see section 4.2). 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). Given that our coverage of Russian municipalities equals or approximates the universe, we do not apply sampling weights.

We also replicate our analysis using Propensity Score Matching (PSM). However, we find that in our setting PSM – unlike Mahalanobis matching – is inadequate for guaranteeing that matches are close in geographical space.[19] The ATT estimates obtained via PSM (available on request) are usually larger than those obtained in the Mahalanobis case, which we consider more conservative.[20]

4.2 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*;[21] in total, our dataset includes 2,333 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).

4.2.1 Science cities database

The science cities database is based on various publicly available sources. An official, definitive list of science cities does not exist, but most of the 95 middle-sized urban centres on our list appear in Aguirrechu (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

[19] Furthermore, while we are able to obtain a similar degree of covariate balance across the two methods, we observe that unlike Mahalanobis, covariate balance is not robust to the replication of the PSM algorithm on a sub-sample of the data. Arguably, the reason is that none among our covariates is a strong predictor of science city status, as the selection criteria for science cities' location were quite diverse.

[20] 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 the quality of matching, we calculate Mahalanobis distances using the logs of covariates with highly asymmetric empirical distributions. In the case of covariates X_{ick} that can take zero values (such as the historical number of R&D institutes and branches of the State Bank) we use the corresponding quantity $x_{ick} = \log(X_{ick} + 1)$.

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

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 specialisation areas (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.[22]

4.2.2 Municipal-level variables

Socio-economic outcomes. Most outcome variables used in the empirical analysis correspond with the endogenous variables from our theoretical framework. In order to measure differentials in the skill level of local inhabitants, we utilise 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.[23]

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 address 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 the combined R&D-ICT sectors from the Russian Statistical Office (ROSSTAT). Note that ROSSTAT data of any kind – including those about small and medium enterprises (SMEs) and municipal budget 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 (standardised in z-scores) observed by satellites from 1992 through 2011,[24] 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 labour 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

[22] 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.

[23] 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.

[24] 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). Examples of economic studies employing night lights as a proxy for economic activity within geographic units for which no alternative data source is available include Hodler and Raschky (2014) and Storeygard (2016).

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 infrastracture, and similar. All measures are converted to 2010 prices using ROSSTAT's official CPI indices.

Historical characteristics. We collect information on historical socio-economic characteristics that could affect both science city status and current outcomes. To account for differences in city size, we use population data from the first post-Second World War census held in the Soviet Union, conducted in January 1959.[25] Since the 1959 census does not provide a breakdown of the population data by educational achievement at the municipal level, 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 proxy for the pre-existing level of human capital of an urban area.

To control for the existing level of industrial development in a municipality, we use two pieces of information. The first is the number of Soviet defence industry plants (factories, research and design establishments) located in each municipality 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, and the geographical dispersion of its branches can be seen as indicative of an area's importance for the Soviet development strategies; see also Bircan and De Haas (2017). Moreover, most science cities needed access to good transportation links, while others had to be located in remote areas far from espionage threats. To account for both factors, we use GIS data to measure municipalities' distance from Russian railroads in 1943[26] and from the post-Second World War USSR borders.

Geographical characteristics. We collect or calculate information about several geographical characteristics of Russian municipalities: 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 specialisations, we collect data on each municipality's access to the coast, a major river or lake.[27]

[25] We would prefer to use population data from the 1940s, but there was no census conducted until 1959; moreover, Second World War affected Russia's demography so much that any figures collected before 1941 are inadequate.

[26] 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 industrialisation. Using information about the railroad network in 1943 is preferable to later dates, because Soviet rail transport became one of the most developed in the world after the Second World War, driven by the country's need to extract – and transport – its natural resources.

[27] 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: descriptive statistics

	Science cities		Other municipalities		
	Obs.	Mean (SE)	Obs.	Mean (SE)	p-value
Latitude	88	55.664 (0.391)	2,250	53.981 (0.108)	0.000
Longitude	88	49.771 (2.387)	2,250	59.955 (0.620)	0.000
January mean °C	88	-11.632 (0.410)	2,250	-13.559 (0.149)	0.000
July mean °C	88	18.535 (0.181)	2,250	18.755 (0.056)	0.247
Average altitude	88	0.169 (0.010)	2,250	0.267 (0.007)	0.000
Minimum distance from railroad	88	0.007 (0.001)	2,250	0.078 (0.005)	0.000
Minimum distance from river	88	0.032 (0.004)	2,250	0.056 (0.001)	0.000
Minimum distance from lake	88	0.118 (0.009)	2,250	0.172 (0.003)	0.000
Minimum distance from coast	88	0.725 (0.044)	2,250	0.730 (0.010)	0.917
Minimum distance from USSR border	88	0.665 (0.037)	2,250	0.679 (0.009)	0.723
Population in 1959	88	67.583 (12.516)	2,250	49.573 (3.242)	0.167
Number of universities in 1959	88	0.557 (0.224)	2,250	0.196 (0.046)	0.132
Number of State Bank branches in 1946	88	1.096 (0.987)	2,250	0.739 (0.977)	0.000
Number of plants in 1947	88	6.205 (1.458)	2,250	2.484 (0.697)	0.023
Number of R&D institutes in 1947	88	0.807 (0.253)	2,250	0.412 (0.222)	0.242
Area in km^2	88	0.692 (0.116)	2,250	7.108 (0.637)	0.000
Current socio-economic outcomes					
Population in 2010	88	131.557 (21.169)	2,250	58.324 (5.871)	0.001
Graduate share in 2010	88	0.225 (0.008)	2,250	0.110 (0.001)	0.000
Postgraduate share in 2010	88	0.006 (0.000)	2,250	0.003 (0.000)	0.000
Fractional patents, 2006-15	88	13.909 (3.489)	2,250	2.265 (1.210)	0.002
Avg. fractional patents, 2006-15	88	0.761 (2.944)	2,250	0.028 (0.107)	0.000
Night lights, 2009-11 average of z-scores	88	1.586 (2.124)	2,250	-0.062 (0.272)	0.000
Avg. salary in R&D and ICT in 2010 (thousands)	71	24.266 (10.001)	2,173	15.366 (7.978)	0.000
Employment share in R&D and ICT in 2010	71	0.036 (6.937)	2,173	0.007 (12.394)	0.000
Number of SMEs in 2010 (thousands, all)	69	3.240 (0.743)	2,140	1.190 (0.067)	0.008
Number of SMEs in 2010 (thousands, manufacturing)	69	0.395 (0.103)	2,038	0.120 (0.008)	0.010
SME labour productivity (all)	69	1.644 (0.085)	2,153	0.794 (0.009)	0.000
SME labour productivity (manufacturing)	67	1.438 (0.085)	2,014	0.768 (0.021)	0.000

Source: Table B.1, municipal-level data described in section 4.2 and authors' calculations.

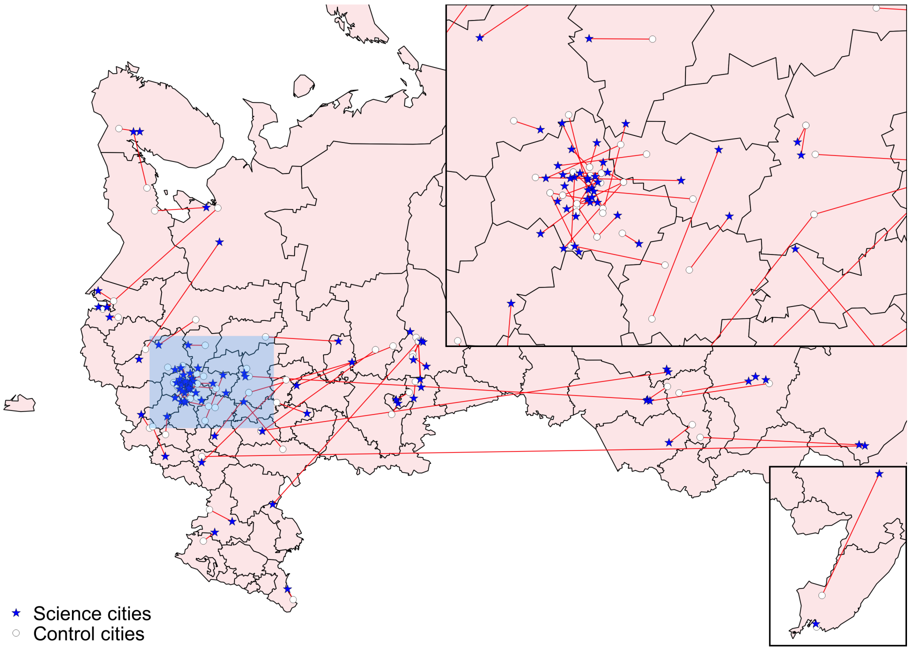
4.2.3 Summary statistics

Table 1 displays the summary statistics for municipal-level characteristics and socio-economic outcomes, distinguishing between municipalities hosting science cities and all other ordinary municipalities. It shows that, on average, science cities were located in more populous and warmer places, with a higher historical concentration of industrial plants, universities, and R&D institutes. In addition, the differences between science cities and other municipalities in the means of all socio-economic outcomes are positive and statistically significant.

4.3 Empirical results

In what follows, we present the results of the municipal-level empirical analysis, starting from the description of the matched sample.

Figure 2: Science cities and their matches



Source: Table B.1, municipal-level data described in section 4.2 and authors' calculations.

4.3.1 Quality of matching

The main matching sample is constituted by 83 municipalities that include a science city and by 65 matched municipalities which do not host any science city. Out of 88 science city municipalities in our dataset, 5 are not matched to any control observations.[28] However, most control observations are matched to, at most, two science cities (a few more in a couple of

[28]These are the closed science cities of Lesnoy, Seversk, Solnechny, Zelenogorsk and Zheleznogorsk.

cases). Figure 2 displays the matched pairs on the map of Russia. Thanks to our choice of covariates, science cities and their counterparts are matched – with a few exceptions – relatively 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 2 displays the standardised mean difference and the variance ratio between treated and control observations, both in the original and in the matched samples. The table shows that matching achieves a remarkable degree of balance in both the first and the second moment, despite the rigidity of the Mahalanobis algorithm and the exact matching requirements that we have imposed on it.

Table 2: Covariate balance: Mahalanobis matching, all science cities

	Standardised bias		Variance ratio	
	Raw	Matched	Raw	Matched
Latitude	0.3592	-0.0423	0.5429	0.9612
Longitude	-0.4503	0.0704	0.5346	1.2422
January mean °C	0.3916	-0.0230	0.2750	1.2278
July mean °C	-0.0854	0.0832	0.4189	1.1847
Average altitude	-0.4050	-0.1003	0.0858	0.9157
Population in 1959	0.1593	0.0165	0.6062	0.9644
Population within 200km in 1959	1.0974	-0.0092	3.6322	1.0543
No. of plants in 1947	0.1623	0.0304	0.2194	1.1487
No. of plants within 200km in 1947	1.1425	0.0176	5.1162	1.0010
No. of universities in 1959	0.1802	0.0511	0.9630	1.0889
(Log) area in km^2	-1.1775	-0.0749	1.1944	0.8325
(Log) no. of R&D institutes in 1947	0.7263	0.0799	4.8844	1.1033
(Log) no. of State Bank branches in 1946	-0.4962	-0.0816	1.5453	1.1959
(Log) dist. from railroads in 1943	-0.8623	0.0110	0.4435	1.0214
(Log) dist. from coastline	-0.0537	-0.0422	1.3513	1.2271
Dist. from USSR border in 1946-91	-0.0359	-0.0539	0.7059	1.0733

Source: Table B.1, municipal-level data described in section 4.2 and authors' calculations.

Note: For each variable in the left column, the table reports both the difference in the variance-standardised mean and the variance ratio between treated and control observations, for both the raw sample and the matched sample. The matched sample is obtained by standard Mahalanobis matching on the variables above, forcing exact matching on: closed city status, and presence of a lake *or* a river in the municipal territory. The number of R&D institutes and branches of the USSR State Bank is increased by one before applying the logarithmic transformation. Matching is one-to-one with replacement. Dist. - distance.

4.3.2 ATT estimation: all science cities

The main estimates of the ATT for our 12 outcomes of interest are reported in Table 3. In what follows, we summarise 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 bias-adjusted estimates for current population).

Among the innovation outcomes, the absolute fractional patents measure estimate is positive and statistically significant (at the 1 per cent level), similarly as the corresponding average measure (significant at the 5 per cent level). These results indicate that between 2006 and 2015, science cities have applied to the EPO, on average, for 9-10 more fractional patents than their matched municipalities, or about 0.7 more fractional patents for each individual with a postgraduate degree.[29] 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 higher by about 6,000 roubles (roughly US$200) in 2010 prices.

Lastly, we examine our proxies of overall economic activity. The average of standardised night lights indicators for 2009-11 registers a statistically significant difference in favour of science cities, although it appears much smaller than the raw difference and, in the bias-adjusted case, it is only significant at the 10 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 characterised 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 weakly significant, if at all, when adjusting for the matching bias. Similar results, which are not displayed in Table 3 for brevity, are obtained for measures of SME density (number of SMEs divided by municipal population). The SME labour 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 an anticipation of our later discussion, we argue that the results seem to point to an economic effect of science cities that 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 $q = 1, \dots, Q$. The size of the simulated unobserved factor is given by parameter $\Gamma \geq 1$, which measures the hypothesised odds of receiving the treatment ($\Gamma = 1$ is the experimental benchmark). In Table 3 we report, for each outcome variable, the lower 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

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

Table 3: Municipal-level results: Mahalanobis matching, all science cities

	Whole sample	*Matched sample* (1 nearest neighbour)				
Outcome	Raw difference	*T*	*C*	ATT	ATT b.a.	Γ* (α = .05)
Population	73.233*** (21.861)	83	66	33.060*** (10.761)	20.834** (9.497)	2.20
Graduate share	0.115*** (0.008)	83	66	0.047*** (0.008)	0.042*** (0.009)	3.45
Postgraduate share	0.003*** (0.000)	83	66	0.002*** (0.000)	0.002*** (0.000)	2.40
Fractional patents	11.644*** (3.676)	83	66	9.828*** (3.183)	9.527*** (3.259)	3.95
Avg. fractional patents	0.733** (0.312)	83	66	0.699** (0.328)	0.675** (0.329)	4.50
Employment share in R&D, ICT	0.028*** (0.004)	66	56	0.019*** (0.003)	0.018*** (0.003)	4.20
Avg. salary in R&D, ICT	8.900*** (1.204)	66	56	5.928*** (1.839)	6.470*** (1.862)	1.85
Night lights (2009-11)	1.649** (0.153)	83	66	0.371*** (0.137)	0.254* (0.132)	3.15
No. SMEs, thousands (All)	2.050*** (0.741)	64	56	0.806** (0.331)	0.376 (0.299)	1.50
No. SMEs, thousands (Manuf.)	0.276*** (0.103)	64	56	0.161** (0.065)	0.116* (0.060)	1.60
SME labour product. (All)	0.850*** (0.084)	64	56	0.305*** (0.108)	0.292*** (0.112)	1.65
SME labour product. (Manuf.)	0.670*** (0.087)	62	55	0.130 (0.120)	0.139 (0.126)	1.00

Source: Table B.1, municipal-level data described in section 4.2 and authors' calculations.
Note: *, ** 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 $\alpha = .05$ type I error. A higher value of Γ is associated with a stronger simulated unobserved factor which affects both the outcome and the probability of receiving the treatment. Full results of the sensitivity analysis are available upon request. Avg. - average; Manuf. - manufacturing; product. - productivity.

treated and control observations.[30] 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

[30]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.

graduate share. They are satisfactorily high (around 2) for the other demographic outcomes and the high-tech salary measure.[31] 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 1.50 and 1.65 – for our SME measures, except for the manufacturing labour 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.

We interpret our results in light of our analytical framework presented in section 3. In the model, highly skilled population and employment in high-tech sectors can be driven by a mechanism of long-run persistence, which traces its roots to the Soviet-era allocation of workers across different cities. For example, highly skilled workers might prefer to live in science cities because they consider them their home, because moving is costly, or because science cities are inherently preferable. Agglomeration forces such as localised knowledge spillovers can reinforce and complement such factors. However, productivity and wages can only be higher in science cities because of agglomeration forces, or due to some other exogenous factors that the model does not account for. Since we trace the differential evolution of 63-83 pairs of matched cities over 20 years following the dissolution of the USSR, we are not inclined to believe that exogenous shocks alone can drive the results that we observe for average patent production, wages in high-tech sectors, and SME labour productivity. Conversely, we interpret this evidence as favourable to the existence of increasing returns to the co-location of highly skilled workers ($\theta_h > 0$), which possibly spills over to lesser skilled workers as well ($\theta_\ell > 0$), as hinted in particular by the SME labour productivity results.

It must be mentioned that while our results are based on one-to-one matching, main qualitative conclusions are not altered in the case of one-to-many matching. In fact, increasing the number of matched nearest neighbours 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 neighbours up to five; however, we do not present these results here due to space limitations.

4.3.3 ATT estimation: historical science cities

Our interpretation of the estimated long-run consequences of science cities, which rests on the interaction between persistence and agglomeration forces, 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. Within our analytical framework, this is identical to the case where the random shocks $\tilde{x}_h$, and possibly $\tilde{x}_\ell$, have a non-zero mean. 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,

[31] $\Gamma = 2$ 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.

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 programme in recent years, although with a less military and more civil focus. We call the remaining science cities "historical".

Table 4: Municipal-level results: Mahalanobis matching, historical science cities

	Whole sample	*Matched sample* (1 nearest neighbour)				
Outcome	Raw difference	T	C	ATT	ATT b.a.	Γ^* (α = .05)
Population	82.854*** (25.398)	69	57	36.803*** (12.147)	23.838** (10.928)	2.30
Graduate share	0.103*** (0.009)	69	57	0.039*** (0.008)	0.034*** (0.009)	2.80
Postgraduate share	0.003*** (0.000)	69	57	0.002*** (0.000)	0.001*** (0.000)	1.80
Fractional patents	7.254*** (2.703)	69	57	4.950*** (1.173)	4.394*** (1.338)	3.35
Avg. fractional patents	0.253*** (0.058)	69	57	0.199*** (0.065)	0.164** (0.065)	3.35
Employment share in R&D, ICT	0.019*** (0.003)	52	49	0.012*** (0.002)	0.012*** (0.002)	3.35
Avg. salary in R&D, ICT	8.481*** (1.403)	52	49	6.701*** (2.003)	7.250*** (2.097)	1.85
Night lights (2009-11)	1.442*** (0.166)	69	57	0.414*** (0.139)	0.265** (0.135)	1.80
No. SMEs, thousands (All)	2.356*** (0.900)	51	48	0.754* (0.396)	0.281 (0.362)	1.35
No. SMEs, thousands (Manuf.)	0.314** (0.125)	51	48	0.168** (0.080)	0.118 (0.075)	1.30
SME labour product. (All)	0.850*** (0.084)	51	48	0.303*** (0.088)	0.311*** (0.101)	1.60
SME labour product. (Manuf.)	0.671*** (0.086)	49	46	0.138 (0.107)	0.153 (0.122)	1.00

Source: Table B.1, municipal-level data described in section 4.2 and authors' calculations.
Note: See the note accompanying Table 3.

The results in Table 4, 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 3. Statistical inferences and sensitivity analyses *à la* Rosenbaum generally confirm this assessment.[32] The removal of *naukogrady* results in a substantial change of the estimated effects only for the patent outcomes. The fractional patent count ATT estimate is about half of the initial estimates in Table 3; for the average fractional patent measure, it is about 70 per cent smaller. Nevertheless, both estimates remain significant at the 1 per cent level and

[32]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.

robust, as evidenced by a Γ^* from the sensitivity analysis well above 3. Notably, the estimates for the employment share and salary in R&D and ICT remain similar to the previous ones.

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 programme, the Russian government has chosen the best former science cities for the newer *naukograd*y programme. In either case, we keep observing a positive differential in favour 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.

4.3.4 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 *naukogrady*, 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 (2015) 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 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 5 summarises the estimates separately for all science cities for which ROSSTAT data are available and for the "historical" subset; 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

Table 5: Municipal-level results: Mahalanobis matching, municipal budgets analysis

	Whole sample	*Matched sample* (1 nearest neighbour)				
Outcome	Raw difference	*T*	*C*	ATT	ATT b.a.	Γ^* ($\alpha = .05$)
Panel A: All science cities						
Total revenues p.c.	-5.290*** (1.312)	63	56	-0.342 (1.473)	0.827 (1.502)	1.00
All transfers p.c.	-8.933*** (0.848)	63	56	-1.667** (0.829)	-1.040 (0.860)	1.05
Tax income p.c.	3.643*** (0.691)	63	56	1.325 (0.963)	1.866* (0.978)	1.50
Total expenditures p.c.	-5.164*** (1.296)	63	56	-0.334 (1.521)	0.902 (1.552)	1.00
Expenditures on education p.c.	-1.355*** (0.488)	63	56	0.511 (0.447)	0.935* (0.497)	1.05
Expenditures on roads p.c.	-0.013 (0.076)	63	56	0.220*** (0.080)	0.270*** (0.083)	1.20
Panel B: Historical science cities						
Total revenues p.c.	-5.943*** (1.330)	50	48	-0.015 (1.541)	1.215 (1.650)	1.00
All transfers p.c.	-9.049*** (0.874)	50	48	-1.657* (0.891)	-1.122 (0.914)	1.00
Tax income p.c.	3.106*** (0.715)	50	48	1.642* (0.900)	2.336** (0.999)	1.50
Total expenditures p.c.	-5.812*** (1.312)	50	48	0.003 (1.585)	1.250 (1.690)	1.00
Expenditure on education p.c.	-1.615*** (2.994)	50	48	0.399 (0.510)	0.991 (0.629)	1.00
Expenditure on roads p.c.	0.002 (0.089)	50	48	0.250*** (0.088)	0.298*** (0.090)	1.25

Source: Table B.1, municipal-level data described in section 4.2 and authors' calculations.
Note: See the note accompanying Table 3. p.c. - per capita.

obtain higher earnings from local taxes. By controlling for historically observable characteristics, 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, although not statistically significant when adjusting for the matching bias. The values of Γ^* are equal or close to 1 for all these outcomes, except for local average tax income in which case $\Gamma^* = 1.5$ – 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 favour. Governmental support to science cities may also exist in the form of direct expenditures appearing only in the federal budget: however, 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 von Ehrlich and Seidel (2015) 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. We only report results on the "education" and "roads" entries, since we never observe 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 do seem to spend on average more on education, this difference is (weakly) statistically significant only when 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 centralised; municipal expenditures in education are mostly related to the maintenance of the local schools. Therefore, it is unlikely that the educational channel can explain the advantage of science cities.

We obtain more interesting results when looking at the "roads" budget entry – which in reality covers a wider range of transportation and physical infrastructure, corresponding more closely with the original von Ehrlich and Seidel (2015) hypothesis. The estimated ATT for per capita expenditures on "roads" is positive and statistically significant at the 1 per cent level in both sub-samples, whether we apply the bias adjustment or not. However, the critical Γ^* value from the sensitivity analysis is very low in both cases, suggesting that this result may be driven by a smaller subset of science cities. For geographical and institutional reasons, the state and maintenance of roads and other terrestrial transportation links is a heartfelt issue in Russia; it is possible that the USSR endowed science cities with better infrastructure of this kind, and *some* post-transition science cities might have played this factor to their advantage in order to keep an edge in the new market economy. While the mechanism postulated by von Ehrlich and Seidel (2015) is certainly worthy of additional exploration, we only have limited evidence to suggest that this is a key factor of science cities' success. Conservatively, we maintain the interplay between persistence and agglomeration forces, as expressed in our analytical framework, as our preferred explanation of the main results.

4.3.5 ATT estimation: demographic dynamics

Another concern is that our results may not be long-lasting. Our model analyses the spatial equilibrium that would emerge in a static context, if workers initially allocated across space by a central planner were suddenly allowed to move. In the real world, however, workers are slowly replaced by younger workers from newer generations. In our framework, persistence forces interacting with spillover effects are modelled as differential preferences between static sets of workers. 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-Second World War 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 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 (1991-92), 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.

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 6 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. All estimates are uniformly smaller, but still statistically significant, if current *naukogrady* are removed from the sample. 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;[33] and are not sensitive to the choice of the threshold. Thus, this analysis provides little evidence in favour 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.

4.3.6 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 standardised 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

[33] 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).

Table 6: Municipal-level results: Mahalanobis matching, "dynamic" analysis

	Whole sample	*Matched sample* (1 nearest neighbour)				
Outcome	Raw difference	T	C	ATT	ATT b.a.	Γ^* ($\alpha = .05$)
Panel A: All science cities						
Graduate share: born ≤ 1965	0.125*** (0.010)	83	66	0.058*** (0.009)	0.051*** (0.010)	4.10
Graduate share: born > 1965	0.109*** (0.007)	83	66	0.035*** (0.008)	0.031*** (0.009)	2.15
Postgraduate share: born ≤ 1955	0.004*** (0.001)	83	66	0.003*** (0.001)	0.003*** (0.001)	2.15
Postgraduate share: born > 1955	0.003*** (0.000)	83	66	0.002*** (0.001)	0.002*** (0.001)	1.85
Panel B: Historical science cities						
Graduate share: born ≤ 1965	0.110*** (0.010)	69	57	0.045*** (0.009)	0.037*** (0.011)	3.10
Graduate share: born > 1965	0.100*** (0.008)	69	57	0.030*** (0.008)	0.027*** (0.009)	1.95
Postgraduate share: born ≤ 1955	0.003*** (0.000)	69	57	0.002*** (0.001)	0.002*** (0.001)	1.65
Postgraduate share: born > 1955	0.003*** (0.000)	69	57	0.002*** (0.001)	0.002** (0.001)	1.45

Source: Table B.1, municipal-level data described in section 4.2 and authors' calculations.
Note: See the note accompanying Table 3.

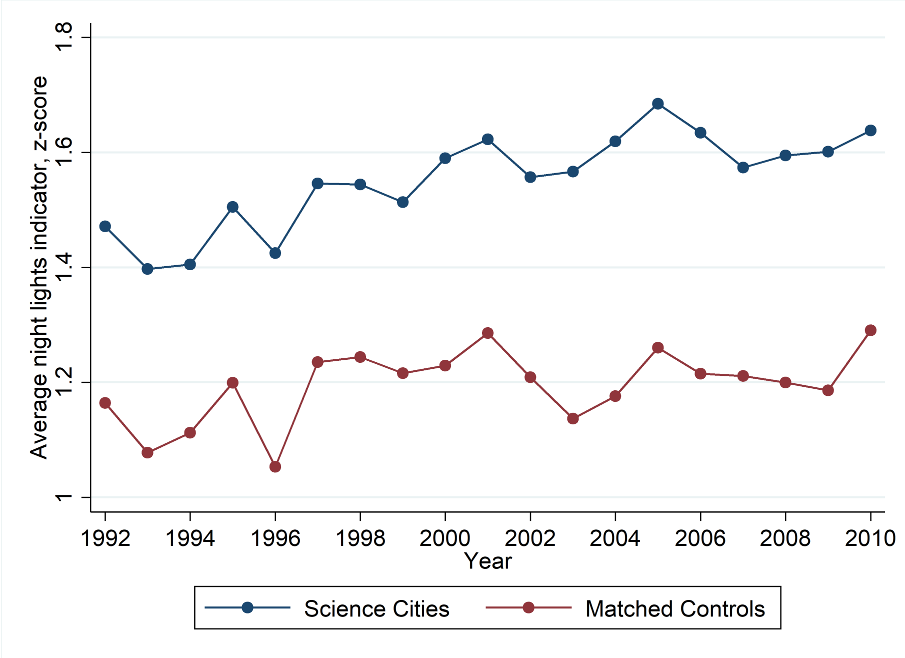
could also be due to – albeit unlikely – an extreme path dependence of some random unobserved factors that are not explained by science city status.

To clear this concern, we use the longitudinal night light data to estimate the following simple regression model:

$$y_{it} = \beta_0 + \beta_1 S_i + \tau_t + \varepsilon_{it} \tag{11}$$

where y_{it} is some night lights measure, S_i is a science city dummy, τ_t is a year effect and ε_{it} is an error term allowed to be autocorrelated in time. Here, parameter β_1 represents the causal effect of science city status, and identification is ensured by restricting the analysis to the matched sample. We expect the estimate(s) of β_1 to be positive, but not necessarily statistically significant when allowing for autocorrelated disturbances. In fact, we have observed that by estimating the ATT separately for each year, the results are often not statistically significant due

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



Source: Table B.1, municipal-level data described in section 4.2 and authors' calculations.

to the statistical dispersion between the two groups.

Table 7: Estimates of parameter β_1 from model (11)

	(1)	(2)	(3)	(4)
Science city (β_1)	0.3573**	0.4754***	0.4146**	0.5306***
	(0.1578)	(0.1680)	(0.1701)	(0.1823)
Year effects	YES	YES	YES	YES
Bias-adjusted estimate	NO	YES	NO	YES
Historical science cities only	NO	NO	YES	YES
Number of observations	3,154	3,154	2,622	2,622

Source: Table B.1, municipal-level data described in section 4.2 and authors' calculations.
Note: *, ** 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.

We report our estimates in Table 7; we calculate standard errors with the Newey-West HAC formula by allowing autocorrelation of ε_{it} up to 10 years (we obtain virtually the same results with longer time windows). 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 3-4, and is statistically significant at the 5 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}, letting the linear correction model vary by year. This results in an even higher measure of β_1 – up to half the measure's standard deviation – which is significant at the 1 per cent level. By restricting the analysis to historical science cities, we obtain even larger estimates, which are reported in columns 3 and 4 (respectively, without and with bias adjustment). To summarise, by looking at night lights as a proxy of economic activity, it appears that the effect of science cities is statistically robust and constant over time, once again offering little support to the mean-reversion hypothesis.

5 Spillovers at the firm level

In this section we assess whether science cities have any effects on the innovation and performance of firms. Like section 4, it is split into three parts: methodology, data and results.

5.1 Methodology

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

$$I^*_{fr} = \beta_0 + \sum_{d=1}^{D} \beta_d W_{fr,d} + \gamma \underbrace{\sum_{s=1}^{S} \exp\left[-\lambda \cdot \text{dist}\left(f,s\right)\right] H_s}_{\equiv G_{fr} = G_{fr}(H_1,\ldots,H_S;\lambda)} + \eta_r + \varepsilon_{fr} \tag{12}$$

where $f = 1,\ldots,F$ indexes firms; $s = 1,\ldots,S$ denotes science cities; r is a subscript for Russian regions; I^*_{fr} is the latent variable associated with one specific innovation binary outcome I_{fr}; $\text{dist}\left(f,s\right)$ is the geodesic distance between firm f and science city s; $\left(W_{fr,1},\ldots,W_{fr,D}\right)$ are D controls available in the data (see section 5.2); H_s is a relevant characteristic of science city s; 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, so we use the following OLS model:

$$\log P_{fr} = \tilde{\beta}_0 + \sum_{d=1}^{D} \tilde{\beta}_d W_{fr,d} + \tilde{\gamma} G_{fr} + \tilde{\eta}_r + \upsilon_{fr} \tag{13}$$

which is analogous to (12). In model (13), P_{fr} represents performance indicators such as the firm's operating revenue (sales) or labour productivity. Functional forms that involve a term akin to G_{fr} are routinely adopted in studies of R&D spillovers (Lychagin et al., 2016) or of agglomeration effects between firms (Drucker, 2012).

The main parameters of interest are γ in the probit regressions and $\tilde{\gamma}$ in the linear regressions. They measure the relationship between the innovation and performance, respectively, of firm f and the characteristic H_s of all science cities s, weighted by the relative geographic proximity between f and each s. To interpret the empirical model more easily, observe that $\exp\left[-\lambda \cdot \text{dist}\left(f,s\right)\right]$ is the exponential decay of a science city's "influence" in space: it is equal to 1 if a firm is located right in the centre of a science city, and it is negligible unless firm f and city s are relatively close. Thus, if a firm is located in a relatively isolated science city, the quantity $\gamma \cdot \hat{\phi}_f$ – where $\hat{\phi}_f$ is the standard normal density function evaluated at the parameter estimates and at firm f's values of the right-hand side variables – approximates the marginal effect of the characteristic H_s of science city s on the probability of a positive realisation of I_{fr} for firm f. Similarly, in linear models $\tilde{\gamma}$ is more easily interpreted as the average change in P_{fr} for firms that are located in a "relatively isolated" science city with characteristic H_s.

These specifications are flexible and vary with the choice of H_s and parameter λ.[34] For both linear and non-linear models, we analyse the dependence of our outcomes of interest with different “agglomeration measures” based on four different science city characteristics H_s that likely relate to its innovation potential. These are: the fractional patents produced in science city s, the graduate share of its population, its postgraduate share and the share of R&D employment. Descriptive statistics and cross-correlations for the resulting firm-level agglomeration measures G_{fr} are displayed in Tables 8 and 9, respectively. We report estimations using fractional patents and postgraduate share-based measures in the same regression.[35]

Table 8: Agglomeration potential measures ($\lambda = 1$): descriptives

	Mean	Std. dev.
Fractional patents	0.0449	0.5641
Graduate share (%)	0.0621	0.6970
Postgraduate share (%)	0.0018	0.0227
R&D employment share (%)	0.0116	0.1379

Source: Table B.1, BEEPS V Russia, municipal-level data described in section 4.2 and authors' calculations.

Table 9: Agglomeration potential measures ($\lambda = 1$): correlations

	Fractional patents	Graduate share (%)	Postgraduate share (%)
Graduate share (%)	0.642		
Postgraduate share (%)	0.547	0.945	
R&D employment share (%)	0.724	0.943	0.843

Source: Table B.1, BEEPS V Russia, municipal-level data described in section 4.2 and authors' calculations.

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. 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,

[34] We report estimates using $\lambda = 1$; the findings are robust to using $\lambda = 2$ or $\lambda = 5$.

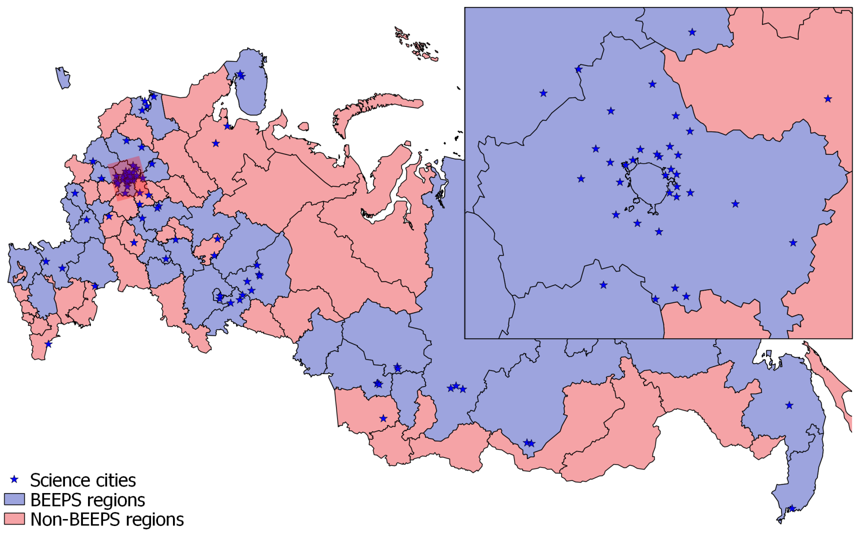
[35] We do not include all of them at the same time because they are highly correlated. Results using one of the other measures in combination with fractional patents are available on request, as are the results with using $\lambda = 2$ or $\lambda = 5$

since we are interested in evaluating in a descriptive sense whether any relationship between science cities and firm-level outcomes extends in space and 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 the Business Environment and Enterprise Performance Survey (BEEPS) merged with Bureau van Dijk's Orbis database, both for Russia only. BEEPS is a firm-level survey conducted by the EBRD and the World Bank, based 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 D.

Figure 4: Location of science cities and regions covered in BEEPS V Russia



Source: 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). Moreover, we are able to match about 75 per cent of BEEPS firms to Bureau van Dijk's Orbis accounting data, giving us access to additional variables from financial statements for a subset of firms; BEEPS V itself 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.

Summary statistics. Table 10 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 10: 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 (employees), Orbis	2,979	3.910	0.062	3.789	4.032
Log (capital), Orbis	3,027	6.169	0.219	5.738	6.599
Log (materials), Orbis	2,936	6.601	0.238	6.132	7.069
Log (permanent, full-time employees), BEEPS	4,211	3.528	0.167	3.200	3.856
Log (operating revenue), Orbis	2,980	6.891	0.217	6.465	7.317
Log (labour productivity), Orbis	2,979	2.956	0.168	2.626	3.286
Log (sales), BEEPS	3,027	17.889	0.209	17.478	18.299
Log (labour productivity), BEEPS	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

Source: BEEPS V Russia and authors' calculations.
Note: Survey-weighted observations (using Stata's svy command). Linearised Taylor standard errors clustered on strata.

5.3 Empirical results

Next, we present estimates of regression models (12) and (13) with $\lambda = 1$; however, we obtain similar results with higher values for this parameter (they are available on request).

5.3.1 *Innovation outcomes*

Table 11 presents the results from the estimation of several probit models with latent variable representation (12) 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 (12), 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 agglomeration potential measure.

Panel A shows that the estimates of fractional patents γ 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 1.8 per cent increase in the probability that the firm engages in R&D, 1.4 per cent increase in the probability that it has introduced a new product in the last three years and 2 per cent increase in the probability that it has ever been granted a patent. The estimates of postgraduate share γ are not significant. These findings indicate that the innovativeness of science cities spills over to the firms that are located sufficiently close to them. 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 manufacturing firms benefit from the spillover effects more than service firms, with R&D and having a patent positively and statistically significantly associated with the patent-based agglomeration potential measure and process innovation positively and statistically significantly associated with the skills-based agglomeration potential measure. The γ coefficient estimates are not statistically significant for service sector firms.

Surprisingly, at a first glance, the estimates of agglomeration potential measure coefficients in Panel C are not significant for firms in high-tech sectors (defined as firms in high-tech, medium-high-tech and high knowledge intensity sectors). They are, however, positive and statistically significant for firms in other sectors in the case of the patent-based agglomeration potential measure; the only exception is process innovation. This could be explained by the relatively small number of high-tech sector firms in our sample. 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 an 11.2 per cent increase in the probability that a young firm has ever been granted a patent, while the same number is only 1.7 per cent for old firms.

Table 11: Firm-level innovation outcomes: probit average marginal effects ($\lambda = 1$)

Agglomeration potential measure	R&D	Product innovation	Process innovation	Has a patent
Panel A: Total sample				
Fractional patents	0.018***	0.014**	0.008	0.020***
	(0.005)	(0.006)	(0.010)	(0.008)
Postgraduate share (%)	-0.167	-0.112	-0.243	-0.123
	(0.183)	(0.187)	(0.358)	(0.186)
Panel B: Allowing different coefficients by sector				
Fractional patents * Manufacturing	0.018**	0.011	-0.006	0.035
	(0.007)	(0.009)	(0.007)	(0.025)
Postgraduate share (%) * Manufacturing	0.001	0.015	0.035***	0.000
	(0.015)	(0.014)	(0.012)	(0.018)
Fractional patents * Services	0.008	0.215	-0.493	0.029
	(0.234)	(0.412)	(0.405)	(0.221)
Postgraduate share (%) * Services	-0.209	-0.484	-0.283	-1.172
	(0.301)	(0.440)	(0.317)	(2.179)
Panel C: Allowing different coefficients by technological level				
Fractional patents * High-tech	-0.054	-0.009	-0.006	0.015
	(0.058)	(0.015)	(0.014)	(0.059)
Postgraduate share (%) * High-tech	0.344	0.492	-0.451	0.201
	(0.395)	(0.848)	(0.340)	(0.365)
Fractional patents * Other	0.022***	0.021**	0.009	0.023**
	(0.006)	(0.009)	(0.011)	(0.010)
Postgraduate share (%) * Other	-0.293	-0.458	-0.187	-0.438
	(0.252)	(0.380)	(0.416)	(0.344)
Panel D: Allowing different coefficients by firm age				
Fractional patents * Young	0.003	0.005	-0.009	0.112**
	(0.012)	(0.018)	(0.013)	(0.053)
Postgraduate share (%) * Young	-1.102	-0.041	-0.917	-0.786
	(0.743)	(0.360)	(0.607)	(0.675)
Fractional patents * Old	0.020***	0.016**	0.009	0.017**
	(0.005)	(0.006)	(0.011)	(0.007)
Postgraduate share (%) * Old	-0.169	-0.114	-0.207	-0.044
	(0.198)	(0.189)	(0.372)	(0.200)
Number of observations	4,040	4,040	4,040	1,863
Number of strata	1,224	1,224	1,224	896

Source: Table B.1, BEEPS V Russia, municipal-level data described in section 4.2 and authors' calculations.

Note: *, ** 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 (using Stata's svy prefix). 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.

To sum up, estimates in Table 11 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 9).

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' labour productivity or sales, by estimating model (13) under different specifications. The results are reported in Table 12; note that for both labour productivity and sales we utilise two different outcome measures, one from BEEPS and the other from Orbis' matched accounting data.[36] BEEPS measures are self-reported during the face-to-face interviews while Orbis data come from the firms' financial reports. Which ones are more accurate is difficult to say. On the one hand, the respondents might disclose only estimates of their sales and employment during the interview. On the other hand, their answers are not used for tax purposes or reported to the authorities, so their answers during the interview may be truthful. Financial reports, in contrast, are used for tax purposes, and firms might underreport their employment and sales. Using measures from both sources might thus provide further insights.

The estimates of $\tilde{\gamma}$ are positive and statistically significant for the patent-based agglomeration potential measure when we use BEEPS-related variables (columns (3) and (4)), and positive and statistically significant for the skills-based agglomeration potential measure when we use Orbis-related variables (columns (1) and (2)). The differences are due to the fact that regressions in columns (1) and (2) control for fixed assets and cost of materials.[37] Once these are taken into account, the patent-spillover effects disappear and skill-spillover effects become important. Having advanced machinery and equipment without having skilled personnel able to use them or without having access to people who can advise on how to use them may mean that they are left idle.

The estimates in Panel B suggest that services firms benefit from the spillover effects more than manufacturing firms, with both operating revenue and labour productivity positively and statistically significantly associated with the skills-based agglomeration potential measure, rather than the patents-based agglomeration measure. A doubling of the index of population with a completed postgraduate education is associated with an almost 350 per cent increase in

[36] Specifically, we employ the measure of operating revenue from Orbis.

[37] We estimate the regressions in columns (3) and (4) of Table 12 using the sample of firms for which both Orbis and BEEPS measures are available. The coefficients estimated using Orbis measures have the same sign and similar significance as the coefficients estimated using BEEPS measures. Results are available on request.

operating revenue and labour productivity of service sector firms located close to science cities. This is a huge effect, but it is important to remember that doubling this index would require a substantial increase in the population with a completed postgraduate education (see Table 8).

Table 12: Firm-level performance outcomes: OLS ($\lambda = 1$)

Agglomeration potential measure	Op. revenue (Orbis)	Labour product. (Orbis)	Sales (BEEPS)	Labour product. (BEEPS)
Panel A: Total sample				
Fractional patents	-0.009	-0.009	0.084***	0.081***
	(0.011)	(0.015)	(0.023)	(0.023)
Postgraduate share (%)	1.089**	1.109**	-1.132	-1.298
	(0.530)	(0.537)	(1.017)	(0.888)
Panel B: Allowing different coefficients by sector				
Fractional patents *	-0.002	-0.011	0.039**	0.038**
Manufacturing	(0.009)	(0.009)	(0.015)	(0.015)
Postgraduate share (%) *	0.217	0.310	-1.784**	-1.817**
Manufacturing	(0.220)	(0.210)	(0.814)	(0.822)
Fractional patents * Services	-0.002	0.003	0.178**	0.165***
	(0.016)	(0.017)	(0.075)	(0.055)
Postgraduate share (%) *	1.495***	1.490***	-1.331	-1.486
Services	(0.390)	(0.374)	(1.236)	(1.112)
Panel C: Allowing different coefficients by technological level				
Fractional patents *	-0.029**	-0.031***	-0.209*	-0.212*
High-tech	(0.012)	(0.012)	(0.121)	(0.126)
Postgraduate share (%) *	0.351*	0.330	-0.423	-0.448
High-tech	(0.199)	(0.204)	(1.276)	(1.315)
Fractional patents * Other	-0.004	0.001	0.084***	0.082***
	(0.015)	(0.016)	(0.024)	(0.023)
Postgraduate share (%) *	1.489***	1.494***	-0.849	-1.048
Other	(0.364)	(0.396)	(1.282)	(1.124)
Panel D: Allowing different coefficients by firm age				
Fractional patents * Young	-0.042*	-0.050**	0.525*	0.589*
	(0.023)	(0.022)	(0.301)	(0.302)
Postgraduate share (%) *	1.367	1.337	-4.769***	-4.792**
Young	(1.118)	(1.147)	(1.849)	(1.872)
Fractional patents * Old	0.002	0.012	0.075***	0.075***
	(0.013)	(0.011)	(0.028)	(0.028)
Postgraduate share (%) * Old	1.042*	0.971	-0.735	-1.137
	(0.608)	(0.620)	(1.700)	(1.405)
Number of observations	2,809	2,809	2,926	2,926
Number of strata	1,086	1,086	1,074	1,074

Source: Table B.1, BEEPS V Russia, municipal-level data described in section 4.2 and authors' calculations.
Note: Simple OLS using survey-weighted observations (using Stata's svy prefix). Orbis measures are based on firm-level data from Bureau Van Dijk's Orbis database, while BEEPS measures are based on firm-level data from BEEPS. Orbis measures use information on the number of employees, fixed assets and cost of materials from Orbis; BEEPS measures use information on the number of employees from BEEPS only, as the other measures are not available for non-manufacturing firms. For other details, see the note accompanying Table 11. Op. revenue - operating revenue; product. - productivity.

A similar effect 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 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.

Young firms located near science cities benefited from patent-based spillovers in terms of having their own patents, but results in Panel D of Table 12 indicate that they had trouble converting this advantage into higher operating revenue and labour productivity. Indeed, both of these outcomes were negatively and statistically significantly associated with the patents-based agglomeration potential measure. Similar to firms in high-tech sectors, young firms may need more time to establish their place in the market.

5.3.3 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, but that the source of these spillover effects differs depending on the outcome. Innovation outcomes, such as R&D, having a patent and product innovation, benefit positively and significantly from the knowledge accumulated in patents produced in science cities. In contrast, performance outcomes, such as operating revenue and labour productivity, benefit positively and significantly from the availability of a highly educated population. The latter is of course correlated with the patenting activity, but it is nevertheless interesting to see that whichever of the two prevails depends on the type of outcome.

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. Service sector as well as 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 analysed 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 programme 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 recognised 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, moreover, 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 today receive resumed support from the Russian government, we conjecture that they are consequent to the interaction between persistence and agglomeration forces, which we illustrate within a simple spatial equilibrium framework. Specifically, 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. By analysing municipal budgets, we rule out alternative explanations that have to do with differential governmental transfers; however, we also obtain some *prima facie* evidence that science cities might have invested more of their resources, per capita, on physical infrastructure. In addition, by examining our data in more detail we find little support for the hypothesis of rapid mean reversion of socio-economic outcomes.

Our contribution extends previous findings about long-run effects of place-based policies to a unique historical programme 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

A. Abadie and G. W. Imbens (2006), “Large sample properties of matching estimators for average treatment effects”, *Econometrica*, *74*(1), 235–267.

A. Abadie and G. W. Imbens (2011), “Bias-corrected matching estimators for average treatment effects”, *Journal of Business and Economic Statistics*, *29*(1), 1–11.

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

J. Ambler, D. J. B. Shaw, and L. Symons (Eds.) (1985), *Soviet and East European transport problems*, London: Croom Helm.

Y. Andrienko and S. Guriev (2004), “Determinants of interregional mobility in Russia: Evidence from panel data”, *Economics of Transition*, *12*(1), 1–27.

D. B. Audretsch and M. P. Feldman (1996), “R&D spillovers and the geography of innovation and production”, *American Economic Review*, *86*(3), 630–640.

C. Beaudry and A. Schiffauerova (2009), “Who’s right, Marshall or Jacobs? The localization versus urbanization debate”, *Research Policy*, *38*(2), 318–337.

Ç. Bircan and R. De Haas (2017), “The limits of lending: Banks and technology adoption across Russia”, mimeo.

N. Bloom, M. Schankerman, and J. Van Reenen (2013), “Identifying technology spillovers and product market rivalry”, *Econometrica*, *81*(4), 1347–1393.

R. Boschma and K. Frenken (2011), “The emerging empirics of evolutionary economic geography”, *Journal of Economic Geography*, *11*(2), 295–307.

R. Bronzini and G. de Blasio (2006), “Evaluating the impact of investment incentives: The case of Italy’s Law 488/1992”, *Journal of Urban Economics*, *60*(2), 327–349.

M. Busso, J. Gregory, and P. M. Kline (2013), “Assessing the incidence and efficiency of a prominent place based policy”, *American Economic Review*, *103*(2), 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.

A. Cheremukhin, M. Golosov, S. Guriev, and A. Tsyvinski (2017), “The industrialization and economic development of Russia through the lens of a neoclassical growth model”, *Review of Economic Studies*, *84*(2), 613–649.

J. M. Cooper (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.

C. Criscuolo, R. Martin, H. Overman, and J. Van Reenen (2012), “The causal effects of an industrial policy”, NBER Working Paper 17842, National Bureau of Economic Research.

D. R. Davis and D. E. Weinstein (2002), “Bones, bombs, and break points: The geography of economic activity”, *American Economic Review*, *92*(5), 1269–1289.

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

K. Dexter and I. 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.

D. Donaldson and A. Storeygard (2016), “The view from above: Applications of satellite data in economics”, *Journal of Economic Perspectives*, *30*(4), 171–198.

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

G. Duranton and D. Puga (2004), “Micro-foundations of urban agglomeration economies”, in J. V. Henderson, and J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, vol. 4, chap. 48, 2063–2117, Cities and Geography.

G. Ellison, E. L. Glaeser, and W. R. Kerr (2010), “What causes industry agglomeration? Evidence from coagglomeration patterns”, *American Economic Review*, *100*(3), 1195–1213.

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

G. Friebel and S. Guriev (2005), “Attaching workers through in-kind payments: Theory and evidence from Russia”, *World Bank Economic Review*, *19*(2), 175–202.

I. Ganguli (2015), “Immigration & ideas: What did Russian scientists ‘bring’ to the US?”, *The Journal of Labor Economics*, *33*(S1), S257–S288, part 2, July.

P. Givord, R. Rathelot, and P. Sillard (2013), “Place-based tax exemptions and displacement effects: An evaluation of the *Zones Franches Urbaines* program”, *Regional Science and Urban Economics*, *43*(1), 151–163.

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

E. L. Glaeser and J. D. Gottlieb (2009), “The wealth of cities: Agglomeration economies and spatial equilibrium in the United States”, *Journal of Economic Literature*, *74*(4), 983–1028.

E. L. Glaeser, H. D. Kallal, J. A. Scheinkman, and A. Schleifer (1992), “Growth in cities”, *Journal of Political Economy*, *100*(6), 1126–1152.

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

L. Gokhberg and L. Mindely (1993), “Soviet R&D resources: Basic characteristics”, in S. Glaziev, and C. M. Schneider (Eds.), *Research and Development Management in the Transition to a Market Economy*, 3–23, Laxenburg, Austria: IIASA.

Y. Gorodnichenko, J. Svejnar, and K. Terrell (2014), “When does FDI have positive spillovers? Evidence from 17 transition market economies”, *Journal of Comparative Economics*, *42*(4), 954–969.

M. Greenstone, R. Hornbeck, and E. Moretti (2010), “Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings”, *Journal of Political Economy*, *118*(3), 536–598.

X. S. Gu and P. R. Rosenbaum (1993), “Comparison of multivariate matching methods: Structures, distances, and algorithms”, *Journal of Computational and Graphical Statistics*, *2*(4), 405–420.

B. H. Hall, J. Mairesse, and P. Mohnen (2010), “Measuring the returns to R&D”, in B. H. Hall, and N. Rosenberg (Eds.), *Handbook of the Economics of Innovation*, Elsevier.

R. Hodler and P. A. Raschky (2014), “Regional favoritism”, *The Quarterly Journal of Economics*, *129*(2), 995–1033.

D. S. Ivanov (2016), “Human capital and knowledge-intensive industries location: Evidence from Soviet legacy in Russia”, *Journal of Economic History*, *76*(3), 736–768.

A. B. Jaffe (1989), “Real effects of academic research”, *American Economic Review*, *79*(5), 957–970.

P. M. Kline and E. 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*, *121*(1), 275–331.

P. M. Kline and E. Moretti (2014b), “People, places, and public policy: Some simple welfare economics of local economic development programs”, *Annual Review of Economics*, *6*, 629–662.

G. M. Lappo 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)*, (1), 20–49, available in Russian only.

S. Lychagin, J. Pinkse, M. E. Slade, and J. Van Reenen (2016), “Spillovers in space: Does geography matter?”, *Journal of Industrial Economics*, *64*(2), 295–335.

A. Marshall (1890), *Principles of economics*, London: Macmillan, 8 edn.

T. Mikhailova (2012), “Where Russians should live: A counterfactual alternative to Soviet location policy”, in MPRA Paper 35938, University Library of Munich, Germany.

E. Moretti (2004), “Workers’ education, spillovers and productivity: Evidence from plant-level production functions”, *American Economic Review*, *94*(3), 656–690.

E. Moretti (2011), “Local labor markets”, in *Handbook of Labor Economics*, Elsevier.

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

E. Moretti, C. Steinwender, and J. 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.

D. Neumark and J. Kolko (2010), “Do enterprise zones create jobs? Evidence from California’s Enterprise Zone Program”, *Journal of Urban Economics*, *68*(1), 1–19.

D. Neumark and H. 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.

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

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

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

K. M. Ruchnov and E. G. Zaitseva (2011), “Crisis of Russian science cities (Krizis naukogradov Rossiyi)”, Tech. rep., mimeo, available in Russian only.

B. G. Saltykov (1997), “The reform of Russian science”, *Nature*, *388*(6637), 16–18.

C. M. Schneider (1994), *Research and development management: From the Soviet Union to Russia*, Heidelberg: Physica-Verlag.

E. Sinani and K. E. Meyer (2004), “Spillovers of technology transfer from FDI: The case of Estonia”, *Journal of Comparative Economics*, *32*(3), 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.

A. Storeygard (2016), “Farther on down the road: Transport costs, trade and urban growth in Sub-Saharan Africa”, *Review of Economic Studies*, *83*(3), 1263–1295.

C. Syverson (2011), “What determines productivity?”, *Journal of Economic Literature*, *49*(2), 326–365.

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

M. von Ehrlich and T. Seidel (2015), “The persistent effects of place-based policy: Evidence from the West-German Zonenrandgebiet”, Working Paper 5373, CESifo.

J. Wang (2013), “The economic impact of Special Economic Zones: Evidence from Chinese municipalities”, *Journal of Development Economics*, *101*(2), 133–147.

Z. Zhao (2004), “Using matching to estimate treatment effects: Data requirements, matching metrics, and Monte Carlo evidence”, *The Review of Economics and Statistics*, *86*(1), 91–107.

Appendices

A An outline of the Soviet innovation system

Throughout its history, the Soviet Union paid significant attention to R&D. Employment in this sector grew from 35,000 in 1922 to more than 2.8 million in 1990, with intensive investment made in R&D facilities and equipment along the way (Gokhberg, 1997). R&D expenditures were also increasing in absolute terms until 1990, when they accounted for about 3.5 per cent of GNP (Gokhberg and Mindely, 1993).

Despite these efforts, the specific characteristics of the Soviet economic and political system harmed the efficiency of the R&D sector. In the absence of market forces, R&D plans and findings were set as a result of bargaining among various parties in the government and the R&D sector while the prices of R&D products were based on their estimated costs. Following the branch-based structure of production and widespread concerns for secrecy, the Soviet scientific sector was dominated by large multipurpose R&D institutions able to exercise monopolistic power in the relevant fields (Schneider, 1994). General policies of autarky implemented by the Soviet Union were also applied to science, resulting in the low level of technology and ideas exchange between the country and the international community (Gokhberg, 1997).

These factors led to various inefficiencies in the sector. Deep governmental involvement biased the concentration of Soviet R&D towards politically important areas such as defence and engineering at the expense of fundamental research such as medical and life sciences (Gokhberg and Mindely, 1993). Being subject to very specific incentives, R&D efforts were often focused on obsolete areas and produced low-quality results. In the struggle for funding, R&D institutions were reluctant to share their technologies and were artificially increasing their staff by keeping older cohorts of researchers in employment.[38] As a result of this disproportional increase in R&D personnel, increases in funding were mainly allocated to wages rather than to buying new state-of-the-art equipment (Schneider, 1994). On the top of these problems, low incentives to innovate in the production sector and weak diffusion of innovations across enterprises and industries hampered the utilisation of the suboptimal output of the Soviet R&D.

R&D was largely separated from higher education, with universities becoming almost exclusively training centres. Basic research was instead concentrated in the system of Academy of Sciences and branch academies of agricultural sciences, medical sciences, and education. Applied R&D, on the other hand, was concentrated in the industrial R&D units, which were established by each branch ministry. Enterprise R&D was the least developed, as enterprises had no incentives to introduce new products and processes. Hence, even in cases where the Soviet Union had a leading position in the development of significant innovations, it fell behind others in diffusion of innovations.

These peculiar characteristics of Soviet R&D have to be taken into account while analysing the

[38] For example, in the late 1980s almost 80 per cent of Soviet researchers were aged 50 years or more, compared with 34 per cent in the United States (Schneider, 1994).

development of the sector during the transition period. Quoting Gokhberg (1997), "[o]nly a part of the R&D sector inherited from the Soviet era can and should be preserved" (p. 9). This view was arguably shared by market reformers in the 1990s, when the Russian R&D sector went through a painful adjustment as part of Russia's transformation into a market economy.

B Science cities in the Soviet Union and Russia

Table B.1: Science cities

				Naukograd			Closed city[d]			Priority specialisation 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
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

[a] Based on Aguirrechu (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 specialisation 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**
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
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

[a] Based on Aguirrechu (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 specialisation 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
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
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
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

[a] Based on Aguirrechu (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 specialisation 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**
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
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 Aguirrechu (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 kilometres
	Coordinates of the municipality centre		GPS coordinates of the centre 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 metres (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 1960-1990, assigned to municipalities in QGIS. Average, median, standard deviation, minimum, and maximum.
	Railroad	Vernadsky State Geological Museum and U.S. 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 U.S. 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.

Continued on next page

Table C.1 – Continued from previous page

Data type	Data sub-type	Data source	Description
	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 centre 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 organisations.
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
State Bank	State Bank branches in 1946	Bircan and De Haas (2017), 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 2006-2015, 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 Atmospheric Administration (NOAA) (`http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html`)	Average night lights for 1992-1994 and 2009-2011, 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 (labour productivity), and SME sales per unit of fixed labour, all sectors and manufacturing only. The SME census does not cover ZATOs, so 16 science cities, which are also ZATOs, are not covered.

Continued on next page

Data type	Data sub-type	Data source	Description
Municipal budget	Average municipal budget revenues and expenditures over 2006-2016	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.

Table C.1 – Continued from previous page

[a] CGIAR is a global partnership of research organisations 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 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, labour 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 D.1).

Table D.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***

Source: 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, organisation 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. More than 90 per cent of Innovation Module interviews were completed face-to-face immediately after the main questionnaire; 5.6 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 analyse 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, about two thirds of the self-reported innovations were reclassified, whereby 24 per cent were no longer classified as an innovating firm, while the remaining innovations were reclassified according to their type. 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. In Russia, only 51.9 per cent of companies that said they introduced new products did product innovation and only 59.7 per cent of companies that said they introduced new processes met the definition of process innovation. 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. We are not able to do the same for organisational and marketing innovation, since there were no corresponding questions in the survey.