American Economic Journal: Economic Policy 2022, 14(3): 322–351
https://doi.org/10.1257/pol.20200289

The Long-Run Effects of R&D Place-Based Policies: Evidence from Russian Science Cities†

By Helena Schweiger, Alexander Stepanov, and Paolo Zacchia*

We study the long-run effects of historical place-based R&D policies: the creation of Science Cities in Soviet Russia. We compare current demographic and economic characteristics of Science Cities with those of localities that were similar to them at the time of their establishment. We find that in present-day Russia, Science Cities are more innovative and productive, host more highly skilled workers, and pay them higher salaries. We interpret these findings as the result of the interaction between persistence and agglomeration forces; we rule out explanations related to the differential use of public resources. (*JEL* J31, J44, O33, O38, P25, R23, R58)

The effectiveness of public support for science and research and development (R&D) is a long-standing issue in the economics of innovation. Both direct subsidies and indirect incentives for research and science typically rest on positive externalities (or other types of market failures) that, in the absence of public intervention, cause underinvestment in R&D. Some specific innovation policies, such as the top-down creation of local R&D clusters, are characterized by a geographical dimension. To evaluate the overall impact of such interventions, it is necessary to assess the spatial extent of knowledge spillovers—one of the three forces of spatial agglomeration first identified by Marshall (1920). The debate about localized innovation policies mixes with the one about broader place-based policies: a key question is whether place-based policies can generate self-reinforcing economic effects that persist after their termination, possibly because of agglomeration forces

*Schweiger: European Bank for Reconstruction and Development (EBRD) (email: schweigh@ebrd.com); Stepanov: European Bank for Reconstruction and Development (EBRD) (email: stepanoa@ebrd.com); Zacchia: CERGE-EI, a joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences (email: Paolo.Zacchia@cerge-ei.cz). Lucas Davis was coeditor for this article. We would like to thank Charles Becker, Ekaterina Borisova, Vincenzo Bove, Julian Cooper, Ralph de Haas, Sergei Guriev, Maria Gorban, Mark Harrison, Denis Ivanov, Sergei Izmalkov, Patrick Kline, Olga Kuzmina, Andrei Markevich, Enrico Moretti, Gérard Roland, Mauro Sylos Labini, and Natalya Volchkova for helpful discussions, as well as two anonymous referees, the participants at the SITE Academic Conference: 25 years of transition, MEIDE conference, 2015 PacDev conference, IEA World Congress, Uddevalla Symposium, World Congress of Comparative Economics, XVIII HSE April International Academic Conference, 2017 Barcelona Workshop on Regional and Urban Economics, 2018 Meeting of the American Economic Association, Meeting of the Urban Economics Association, and seminars at the EBRD, Higher School of Economics, New Economic School, UC Berkeley, University of Genoa, University of Cergy-Pontoise, École Polytechnique, University of Nottingham, University of Warwick, and CERGE-EI for their comments and suggestions. Irina Capita, Georgi Chavchanidze, Evgeniya Dubinina, Jan Lukšič, Margarita Pavlova, Arseniy Scherbov, 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.

†Go to https://doi.org/10.1257/pol.20200289 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

at work. Absent any 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; Kline and Moretti 2014b).

Localized innovation policies are a popular policy intervention, but evidence about their effectiveness is scarce even in the short run, let alone the long run. This paper is one of the few analyzing such policies and their long-run impact. We do so by examining the legacy of Russian Science Cities: 95 middle-sized urban centers that were created or developed by the Soviet government in the territory of modern Russia with the purpose of concentrating highly specialized, strategic R&D facilities. These cities were typically shaped around a specific technological purpose; the Soviet government relocated scientists, researchers, and other high-skilled workers from elsewhere in the Soviet Union to work in the newly created establishments. The creation of Science Cities was prompted by the technological and military competition between geopolitical blocks during the Cold War; unsurprisingly, most of them were specialized in military-related fields, such as nuclear physics, ballistics, aerospace, and chemistry. Russia maintains a comparative technological advantage in these sectors to this day.

While one may question whether the institutional context of Russian Science Cities is comparable to that of other industrialized countries, this historical experience stands out in two respects. First, it greatly diminishes selection bias concerns due to unobserved determinants of future development, which typically affect studies about innovative clusters in other countries. The allocation of resources in the Soviet command economy was managed according to suboptimal, often erratic rules of thumb—especially so for highly secretive projects, which were managed by a handful of bureaucrats lacking the advice of experts (Gregory and Harrison 2005; Harrison 2017). Soviet science was, in fact, a largely secretive business managed by the internal police and the army (Siddiqi 2008, 2015). A rule of thumb appears most evident in the choice of Science Cities' locations (Agirrechu 2009), which was based on a secrecy-usability trade-off: the Soviet leaders prioritized places that offered better secrecy and safety from foreign interference (in the form of R&D espionage) or that were otherwise easy to control by governmental agencies, by virtue of geographical proximity. The potential for economic development and local human capital accumulation was arguably not, at the margin, a determinant of location choice.

Second, the transition to a market economy that followed the dissolution of the USSR resulted in a large negative shock to Russian R&D: direct governmental R&D expenditure as a percentage of GDP fell by about 63 percent, causing half of the scientists and researchers of post-1991 Russia to lose their jobs. Consequently, state support for Science Cities was abruptly suspended for obviously exogenous reasons; only in the 2000s was it partially resumed for 14 of the former towns, which today bear the official name of Naukogrady (Russian for "Science Cities"). These historical developments provide us with a unique opportunity to study the long-run consequences of an exogenous spatial reallocation of highly skilled workers, decades after the termination of the program that originally motivated such reallocation. In addition, by analyzing historical Science Cities separately from modern Naukogrady, we are able to evaluate to what extent the modern characteristics of the former depend on the long-run effects due to the Soviet-era policy, rather than on current government support.

These distinctive institutional features motivated us to build a unique, rich dataset covering geographical, historical, and present characteristics of Russian municipalities to answer more general questions about the effects of innovation-centered place-based policies. With it, we estimate the effect of the past establishment of a Science City on present-day municipal-level human capital (measured as the share of the population with either graduate or postgraduate qualifications), innovation (evaluated in terms of patent applications), and various proxies of economic development. To give our estimates a causal interpretation, we match Science Cities to other localities that, at the time of selection, were similar to them in terms of characteristics that could affect both their probability of being chosen and their future outcomes and were on a similar population growth trend. Our main identifying assumption is that, conditional on our matching variables, the choice of a locality was determined at the margin by factors that would be independent from future, posttransition outcomes.

Our findings can be summarized as follows. At present, Soviet-era Science Cities still host a more educated population, are more economically developed, employ a larger number of workers in R&D and ICT-related jobs, and apply for more patents than the localities comparable to Science Cities at the time of the program's inception. In addition, workers in former Science Cities receive substantially higher gross monthly salaries: roughly US$250 per month for high-skilled occupations and US$100–120 per month for low-skilled occupations. The estimated treatment effect is typically lower than the raw sample difference for all outcome variables except those related to patent applications, for which no ex ante bias can be attested from our estimates. The results remain largely unchanged when modern Naukograds are excluded from the analysis, but the point estimates relative to patent filings and the share of workers employed in the R&D or ICT sectors decrease substantially. This suggests that current government support is an important stimulant for employment in R&D and ICT sectors. A more in-depth analysis of demographic outcomes and economic development (proxied by night lights) reveals little to no evidence of mean reversion.

We interpret the findings in light of a spatial equilibrium model, inspired by Glaeser and Gottlieb (2009); Moretti (2011); and Allen and Donaldson (2020), that incorporates both path-dependence and agglomeration forces. We estimate the equilibrium predictions of our model on the matched sample; this reveals large estimates of the parameters that embody path-dependence as well as robust agglomeration elasticities—in a neighborhood of 0.3—stemming from a higher concentration of high-skilled workers. Our model tentatively suggests that the government should provide a subsidy equal to about 150 percent of the local ex ante wage to high-skilled workers in order to reproduce allocations of labor that would mimic that of historical Science Cities achieved under the command economy. This finding contrasts with the study by von Ehrlich and Seidel (2018)—henceforth, vES—of the formerly subsidized West German municipalities that used to border the Iron Curtain. While their empirical analysis rules out agglomeration effects, they propose persistence in public goods investment as the explanation of their measured long-run effects. Our paper provides one of the first assessments of the vES hypothesis through the analysis of municipal budget data. We observe no evidence of differences in the overall expenditure or patterns in the use of resources between Science Cities and their matched counterparts.

Our paper adds to the growing body of studies about the evaluation of place-based policies.[1] It is most directly related, conceptually and methodologically, to the studies by Kline and Moretti (2014a) about the Tennessee Valley Authority; Fan and Zou (2021) on China's "Third Front" state-driven industrialization of inner China; and Heblich et al. (2020) about the "Million Roubles Plants" built in China with Soviet support during the early Cold War. While these contributions also exploit unique historical circumstances of political, geographical, and military kind to uncover long-run effects from historical place-based policies, the Science Cities program stands out, as it specifically concerned investments in knowledge, rather than in physical capital.

We also contribute to the more general search of agglomeration effects—and in particular, localized knowledge spillovers—in urban and regional economics. This has long been a traditional field of investigation for economic geographers, with a particular interest in innovation clusters. Following the seminal contributions by Jaffe (1989); Glaeser et al. (1992); Audretsch and Feldman (1996); and others, a large literature has developed. The issue has caught the attention of economists working in other fields, with several papers focusing on local productivity spillovers, for example, Moretti (2004); Ellison, Glaeser, and Kerr (2010); Greenstone, Hornbeck, and Moretti (2010); and Bloom, Schankerman, and Van Reenen (2013).

This paper's institutional setting relates it to existing research on the consequences of historically massive forms of government intervention on long-run economic and technological development, in Russia or elsewhere. Cheremukhin et al. (2017) argue that the "Big Push" industrialization policy enacted in the USSR under Stalin did not succeed in shifting Russia onto a faster path of economic development. Mikhailova (2012a, b) evaluates negative welfare effects due to the regional demographic policies enacted in the Soviet Union; she also finds that locations hosting Gulag camps grew significantly faster than similar places without camps.[2] In the more specific case of R&D policies, 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, Steinwender and Van Reenen (2021) show that increases in government-funded R&D for military purposes have positive net effects on the TFP of OECD countries, despite crowding out private expenditures in R&D.

This paper is organized as follows. Section I summarizes the history and geographical patterns of Science Cities. Section II outlines the empirical methodology. Section III describes our data. Section IV discusses the empirical results. Section V outlines our spatial equilibrium model and its empirical implications. Lastly, Section VI concludes.

[1] For a recent survey of the empirical research, see Neumark and Simpson (2015).

[2] Toews and Vézina's (2020) study about the legacy of gulag camps that hosted highly skilled political prisoners echoes this result and corroborates many of our findings.

I. Historical and Institutional Background

The former Soviet Union was among the pioneers of public investment in science and in place-based policies focusing on R&D. In the context of the Cold War competition, the Soviet leadership prioritized 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 defense (Cooper 2012). Moreover, science was mostly the army's responsibility (Siddiqi 2015). Science Cities emerged in this environment. We identify 95 middle-sized urban centers that the Soviet government endowed with a high concentration of research and development facilities, each devoted to a particular scientific and technical specialization. Science Cities began to develop around strategically important (military) research centers from the mid-1930s;[3] however, the majority of them were established after World War II, especially in the 1950s.

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

Starting in the 1940s, with the need to protect the secrecy of nuclear weapons in the Cold War environment (Rowland 1996), many Soviet municipalities of military importance were "closed" to external access to maintain security and privacy. Nonresidents 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 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; we take this into account in our empirical analysis.

[3] The model of innovation followed by the Soviet authorities since the early 1930s was the creation of "special-regime enclaves intended to promote innovation" (Cooper 2012). These enclaves first appeared as secret research and development laboratories (so-called Experimental Design Bureaus or sharashki) in the Soviet Gulag labor camp system (Siddiqi 2008).

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 percent to a mere 0.74 percent in 1992, a fact made even more dramatic as Russian GDP shrank by about 50 percent in the early years of the transition. Wages plummeted, and consequently, total employment in R&D fell by about 50 percent.[4] Over the transition, there was little to no recovery from these initial shocks (see online Appendix A for more details). This has inevitably affected Science Cities; while detailed information about their government funding in the 1990s is not available, anecdotal evidence speaks of an effective discontinuation of the military research programs that Science Cities were responsible for, at least until the government reestablished direct support for the 14 modern Naukograds in the early 2000s (though this time without explicit military focus). Our analysis of recent municipal budgets in Section IV confirms that Science Cities today receive, if anything, lower governmental transfers than comparable towns.

The Choice of Locations That Hosted Science Cities.—Since Science Cities were created in secret and in a staggered fashion, typically to respond to perceived military-technological threats coming from the West, there exists no detailed, comprehensive account of how their locations were chosen.[5] To inform our empirical analysis, we draw on the historical meta-analysis by Gregory and Harrison (2005), which examines the archives produced by the Soviet state between its inception and 1960. The authors note that in the USSR, the allocation of resources was based on very imperfect and informal "rules of thumb" rather than efficiency and optimality criteria.[6] This was ultimately caused by a combination of Hayekian information problems and a failure of the Soviet political economy model—at all levels of its bureaucratic apparatus—to credibly commit to a set of contingent, efficient rules for the management of a planned economy. In the case of secretive decisions, these problems were exacerbated by a security-usability trade-off (Harrison 2017). In short, the Soviet leaders commonly faced a dilemma: while sharing secret information and

[4] In the USSR, the wages of scientists were about 10–20 percent above average. They dropped to about 65 percent of the average wage in 1992, following the state's withdrawal 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.

[5] The Russian Ministry of Defense started working on a series of publications documenting the activities of its Soviet predecessor until 1960. However, those available so far only cover the period up to 1941. In the future, these might allow historians and economists to better reconstruct the process.

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

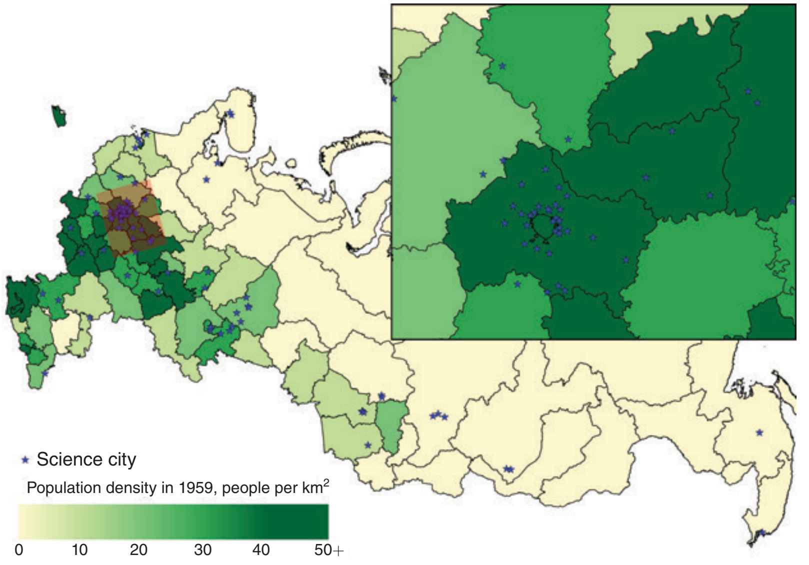

Figure 1. Location of Science Cities and Regional Population Density

Sources: Online Appendix Table B.1 and ROSSTAT.

choices with competent agents might jeopardize the secrecy, not doing so would entail the opposite risk of taking ineffective if not harmful decisions.

The geographical pattern of Science Cities' locations is described in detail by Agirrechu (2009) and depicted in a choropleth map of modern Russian regions, distinguished by their population density (Figure 1). Following Agirrechu (2009, 21), Science Cities can be split in two groups of approximately equal size by type of location. The first group is composed of localities situated in urbanized areas (e.g., in the Moscow region)—they "hosted mainly organizations focusing on theoretical research."[7] By contrast,

> [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 centers, highways, industrial facilities, and production fields. The majority of them were surrounded by forests which served as a natural protection from espionage. In these Science Cities, the core enterprises were military-related R&D institutes, design bureaus, pilot plants, and test sites. Agirrechu (2009, 21)

This pattern can be interpreted as the consequence of a rule of thumb determined by the security-usability trade-off. The specific R&D to be conducted in a prospective Science City would determine the dominant side of the coin; accordingly, Soviet leaders would either choose some remote, highly secretive (but perhaps not

[7] This group included so-called academic towns—semi-isolated neighborhoods of a larger city, endowed with R&D facilities, housing for R&D staff and their families, as well as basic local infrastructure. Because they were part of a larger city, we exclude them from our analysis (see also Section III).

too usable) location[8] or an easy to control (but less secluded) place.[9] Agirrechu (2009) also cites other technological and geographical factors that constrained the choice of location and that, again, depended on the specific R&D specialization: heavy industry and nuclear technology need large amounts of water for their operations, therefore Science Cities focused on those areas were typically built close to rivers or lakes; Science Cities devoted to military shipbuilding and design had to be located on the coast. Science Cities necessitating timely access to production inputs had to be placed closer to transportation links, such as railroads.

The locations chosen to host Science Cities were not random; they typically belonged to the more densely populated and urbanized areas of Russia. The quality of our empirical analysis, however, depends on the extent to which the chosen locations embed unobservable factors that made them more (or less) likely to embark on a path of faster demographic and economic development, relative to other places that were located in the same areas and were otherwise observationally identical to the chosen locations at the time of their selection. It is impossible to provide a definitive response to this question with the currently available archival documentation. Both the historical analysis and the anecdotal evidence, however, clearly point to a negative answer. Harrison (2017) cites a series of shortcomings of the decision process taken by Soviet leaders under secrecy: the relevant information was limited to a handful of trusted bureaucrats, usually selected on the basis of their loyalty to the regime rather than on their competence; expert advice was typically absent; decisions were often taken with information limited by other state secrets. In this context, it appears unlikely that choices taken in a planned economy for military and strategic reasons might have been informed by subtle economic factors.[10] The anecdotal evidence speaks of very idiosyncratic criteria that often determined the exact locations of certain Science Cities; examples include Sarov and Snezhinsk.[11]

II. Empirical Methodology

We compare the long-run outcomes Y_{iq} of municipalities hosting Science Cities with those of other municipalities that were similar in terms of geographical and

[8] This bears similarities to China's "Third Front" industrialization policy that is examined by Fan and Zou (2021). Military considerations also determined the locations of the "Million Roubles Plants" built in China and studied by Heblich et al. (2020).

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

[10] In personal correspondence with us, the Emeritus Professor Mark Harrison concluded that the Soviet leaders would choose a Science City location "[a]s a child with a blindfold would pin the tail on the donkey, I would guess."

[11] These two places provide an 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 the 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 specialized in production of enriched uranium were also located far from each other.

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

The central identifying assumption is motivated by our earlier historical discussion. We assume that the long-run socioeconomic outcomes of both Science Cities and their matched counterparts are orthogonal to the treatment, conditional on relevant geographical and historical characteristics and on the paired locations belonging to the same (type of) geographical region. To lend credibility to our approach, we include a number of indicators (detailed below) about the military, scientific, and economic importance of each municipality in postwar Russia, in addition to matching on cities of equal size and the geographical constraints described by Agirrechu. We take measures to ensure that cities are matched sufficiently close in space, especially in the densely populated parts of Russia such as the Moscow region, but not excessively so, to help ensure that our results are not driven by violations of the Stable Unit Treatment Value Assumption (SUTVA).[12] In particular, we include in the conditioning covariates both geographical coordinates (latitude and longitude) and the density of historical population and factories within a radius of 200 km to control for potentially spatially correlated local economic conditions. Moreover, we restrict the minimal distance between matches to 50 km. Thanks to all these measures, we obtain a matched sample that conforms to our identifying assumptions and allows us to test for possible violation of the SUTVA, as detailed in Section IV (see online Appendix E for additional methodological discussion).

Another concern with our matching approach is that historical characteristics observed at a particular moment in time are not sufficient to account for the dynamic development path of different localities: a Science City might look like a matched control municipality at some point in postwar times even if it was already enjoying faster population and economic growth. If pre-trends are likely to continue, they would threaten our causal evaluation and would motivate approaches in the spirit of the "synthetic control matching" (SCM) method, for example, by matching on each municipality's time series of urban population. Applying SCM to our many outcome variables is unworkable; furthermore, it would not be devoid of shortcomings. As illustrated by Ben-Michael, Feller, and Rothstein (2020), SCM is biased with short panel dimensions; they propose an "augmented" SCM through a covariates-based bias correction or equivalently, the inclusion of both covariates and pretreatment outcomes in a "hybrid" multivariate matching procedure. Inspired by the latter option, we implement a Mahalanobis matching algorithm in which the vector of

[12] Violations of the SUTVA might occur in both ways. On the one hand, the economic effect of Science Cities could extend beyond their borders, so matching two municipalities too close to each other would lead to downward-biased ATT estimates. On the other hand, Science Cities that are too close to matched control cities could drain resources from the latter, leading to upward biases. If the SUTVA does not hold, these two mechanisms could partly offset one another.

covariates includes the observations of all municipalities' urban population at several points in the past.

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

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

where $\mathbf{x}_{ic}$ is the vector of the K observable covariates for municipality i of type $c = s, z$; $\Sigma^{-\frac{1}{2}}$ is the inverted Cholesky decomposition of the empirical variance-covariance matrix of the covariates, Σ, while $\mathbf{W}$ is a matrix of weights obtained via a "genetic" algorithm aimed at optimizing covariate balance (Diamond and Sekhon 2013). Matching is performed with replacement so that a control municipality can be linked to multiple treated cities. We impose exact matching on selected dummy variables;[13] importantly, we match Science Cities subject to the "closed city" status described in Section I to non–Science Cities that experienced similar restrictions (typically, these are places hosting military bases but lacking an R&D content). To improve on balance over the entire distributions of covariates, we calculate Mahalanobis distances taking the logs of continuous covariates with asymmetric empirical distributions.[14]

In our main analysis, we match Science Cities s to control municipalities z one-to-one, conservatively accepting a higher variance for our estimates in exchange for a lower bias. We derive a unique association of treated-control observations that is based on the original set of 84 Science Cities in our dataset. However, most ATT estimates are performed on a subset of this matched sample, either because for some Science Cities the information about 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, perform statistical inference by calculating standard errors based on conventional formulas (Abadie and Imbens 2006, 2011), and adjust p-values using Holm's method to control the family-wise error rate (FWER). Since our coverage of Russian municipalities equals or approximates the universe, we do not apply sampling weights.

III. Data

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;[15] in total, our dataset includes 2,338 such municipalities (the two

[13] This means that we pair municipalities with the smallest Mahalanobis distance conditional on them having the same observed values of these dummy variables.

[14] We calculate logs of $(X_{ick} + 1)$, thus $x_{ick} = \log(X_{ick} + 1)$.

[15] In this paper, we use the English term "municipality" to denote the municipal formations ("municipal'nye obrazovaniya") of Russia, that is, units at the second administrative level (akin to US counties; also called "rayons"). Most urban localities in Russia (places denominated "gorod") are assigned to separate municipalities. We

large cities of Moscow and St. Petersburg are excluded). We used GIS software 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. Additional information, sources, and references are provided in online Appendixes B (for the Science Cities database) and C (for the municipal-level information).

List of Science Cities.—The Science Cities database is based on various publicly available sources. Since Science Cities were established in secret, an official and definitive list does not exist; the extant lists are not exhaustive, having been put together following the dissolution of the USSR. Most of the 95 middle-sized urban centers on our list appear either in Agirrechu (2009) or in the other sources listed in online Appendix B. The database contains information on the location of each Science City, the year the locality was founded, the year in which it became a Science City in the Soviet Union (and the year it became a Naukograd in Russia, where applicable), the type of Science City, whether it was a closed city in the past or is still closed now, and its priority areas of specialization (Table B.1 in the online 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.[16] We exclude four Siberian regional capitals that hosted academic towns from our analysis. While these are "archetypical" Science Cities, the academic towns in question were incorporated in the municipalities of the respective regional capitals, and we are unable to collect statistical information that is suburb-specific. Hence, keeping these municipalities in the treatment or in the control group would contaminate either. Ultimately, we end up with 84 municipalities hosting at least one Science City.

Socioeconomic Outcomes.—In order to measure differentials in the skill level of local inhabitants, we use 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 who completed any form of postgraduate education.[17] We proxy innovation by the total count of local inventor addresses that appear on patent applications filed to the European Patent Office (EPO) with the priority years falling into periods 1978–1999 and 2000–2013. Each address is weighted by the inverse of the number of inventors who appear on the relevant patent application; we call this measure (local) fractional patents.

use the word "region" to refer instead to federal subjects ("oblast," "kray" or "respublika"), that is, units at the first administrative level.

[16] Our list contains four Science Cities for which only their Soviet-era nomenclature is publicly available: Krasnodar-59, Novosibirsk-49, Omsk-5, and Perm-6. Since their exact location is still unclear, we exclude them 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.

[17] 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.

We also collect information from the Russian Statistical Office (ROSSTAT) on total employment and per capita wages in two sectors: R&D, ICT, and related "high-skilled" services sector, and the less skill-intensive wholesale and retail sector.[18] Lastly, since accurate GDP data at the municipal level are unavailable in Russia, we use several proxies for economic activity: the night lights intensity (standardized in z-scores) observed by satellites between 2009 and 2011[19] as well as the density and labor productivity of local SMEs across all sectors of the economy from the 2010 SME census by ROSSTAT.

Municipal Budgets and Local Public Goods.—We obtain data on the budgets of Russian municipalities for 2006–2016 through ROSSTAT. Once again, the information is missing for all closed cities in the sample. On the revenue side, we are able to differentiate between direct revenues (for example, from local taxes) and transfers to the municipalities from both the federal and regional governments.[20] In addition, we are able to distinguish local expenditures by category, such as education, health care, local infrastructure, and similar. All measures are converted to 2010 prices using ROSSTAT's official CPI indexes and averaged over 2006–2016. ROSSTAT also allows us to access data about certain public goods available in Russian municipalities in 2010. Specifically, we calculate the length of local roads that is lit during the night as well as the number of museums, theaters, and libraries in a municipality: services most amenable to the better educated.

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

Historical Characteristics.—We collect information about historical social and economic characteristics that could affect both Science City status and current outcomes. To account for historical differences in city size and population density within a controlled radius, we use population data from the first post-World War II census held in the Soviet Union, conducted in January 1959, which provides figures for all urban and large rural localities of that time.[21] To control for pre-trends in population and economic growth as discussed in Section II, we complement this data with the population of Russian cities in 1897, 1926, and 1939 as reconstructed by

[18] Note that ROSSTAT data sources of any kind are typically never available for closed cities due to national security considerations.

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

[20] These transfers do not include direct government grants to R&D facilities and companies located within a municipality for various purposes, including salaries. Collecting this information across all municipalities in Russia would be close to impossible.

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

Mikhailova (2012a, b) using historical census data from each year. Unfortunately, the 1959 census does not provide a population breakdown by educational achievement at the municipal level; to proxy for the pre-existing level of human capital of an urban area, we use data on the number of higher education institutions located in a municipality in 1959 as well as the number of local R&D institutes in 1947. We also collect population data from the 1989 census for use in the structural analysis in Section V.

To control for the existing level of industrial development in a municipality, we use two pieces of information. The first is the number of the Soviet defense industry plants (factories, research and design establishments) located in each municipality and its surroundings (within 200 km) in 1947. The second is the number of local branches of the State Bank of the USSR in 1946, obtained from its archives. This institution was an instrument of the Soviet economic policy; the geographical dispersion of its branches was indicative of an area's importance for the Soviet development strategies (Bircan and De Haas 2020). To account for the fact that some Science Cities needed access to good transportation links, while others had to be located in remote areas far from espionage threats, we use GIS data to measure municipalities' distance from Russian railroads in 1943[22] and from the post-World War II USSR borders.

Summary Statistics.—Summary statistics for the outcomes, the geographical and historical characteristics illustrate that Science Cities were typically located in more populous and warmer places, with a higher historical concentration of industrial plants, universities, and R&D institutes (see Table D.1 in online Appendix D). Moreover, mean differences between Science Cities and all other ordinary municipalities are positive and statistically significant for most outcomes.

IV. Empirical Results

Our empirical estimates are obtained from a matched sample consisting of 78 municipalities that include a Science City and 59 matched municipalities that do not host any Science City. In online Appendix E, we provide an additional description of the matched sample, including the analysis of covariate balance and a visual representation on the map of Russia. The matched sample is obtained by imposing exact matching on five dummy variables: closed city status, direct access to the coast, access to rivers or lakes, being a mountainous municipality (with an average altitude higher than 1000 km), and has R&D institutes in 1947. Out of 84 Science City municipalities in our dataset, 6 are not matched to any control observations (there are no suitable matches for them under our exact matching constraints), while most control observations are matched to, at most, 2 Science Cities (a few more in a couple of cases).

[22] In the Soviet economy, railroads were the workhorse of the transportation network; road transport played only a secondary role (Ambler, Shaw, and Symons 1985). Most of the railroads' construction took place in czarist Russia.

In what follows, we discuss our main results as well as a number of additional estimates that explore potential drivers of the main effects. Note that our ATT effects estimates can be interpreted as causal effects of Science Cities only if the assumed counterfactual scenario is one where the government had not trained extra scientists as part of the Science Cities program, assuming negligible general equilibrium effects. In a counterfactual scenario where the government had trained extra scientists but decided to distribute them uniformly across similar locations, our control municipalities would no longer be an appropriate counterfactual (we explore alternative counterfactual scenarios with the aid of a model in Section V). At the same time, if the USSR had established a much higher number of Science Cities than it actually did, general equilibrium effects would be much harder to neglect.

ATT Estimation: All Science Cities.—The main estimates of the ATT for our 11 outcomes of interest are reported in Table 1. Science Cities seem to be, on average, more populated than their matched counterparts, by about 26,000–28,000 people. This difference is driven, for the most part, by the more educated segments of the population. The share of inhabitants who attained graduate education is about 6–7 percentage points higher in Science Cities. Similarly, Science Cities still host more people with postgraduate qualification (by about 0.3 percentage points). These estimates are substantially smaller than the corresponding raw differences in the full sample but are generally statistically significant based on statistical tests performed with a FWER adjustment using Holm's method that groups all the outcomes listed in Table 1 together with the "demographic dynamics" outcomes listed in Table 3.[23]

Among the innovation outcomes, the fractional patent applications estimate is positive and statistically significant both before and after 2000. These results indicate that, before 2000, Science Cities have applied to the EPO, on average, for about three additional fractional patents; the number rises to around nine in the period from 2000 onward.[24] 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 sector register differences that are positive and statistically significant (at the 1 percent level). The share of jobs in this sector is 2 percentage points higher in Science Cities, and these jobs pay a RUB7,500 (in 2010 prices; roughly US$250) higher gross monthly salary. Interestingly, we observe a similar but smaller effect on wholesale-retail sector wages, too.

Lastly, we examine our proxies for economic activity. The 2009–2011 average standardized night lights indicator registers a statistically significant difference in favor of Science Cities, although it appears much smaller than the raw difference. The difference amounts to about 45–55 percent of the indicator's standard deviation

[23] The replication package (Schweiger, Stepanov, and Zacchia 2022) calculates the *p*-values of statistical tests performed with a FWER adjustment using Holm's method and reports the stars associated with the standard levels of significance.

[24] We obtain similar results if we use absolute, as opposed to fractional, measures of patent applications or per capita measures.

Table 1—Municipal-Level Matching, Main Analysis: All Science Cities

	Whole sample	Matched sample (1 nearest neighbor)				
Outcome	Raw difference	T	C	ATT	ATT b.a.	Γ^*
Population in 2010 (thousands)	35.524 (9.934)	78	59	27.256 (10.869)	26.367 (9.803)	2.15
Graduate share in 2010	0.113 (0.009)	78	59	0.067 (0.010)	0.059 (0.010)	4.25
Postgraduate share in 2010	0.003 (0.000)	78	59	0.003 (0.001)	0.003 (0.001)	2.80
Fractional patents before 2000	2.610 (1.002)	78	59	3.014 (0.713)	2.822 (0.713)	3.40
Fractional patents since 2000	8.197 (2.596)	78	59	9.260 (2.160)	8.608 (2.170)	3.40
Employment share in R&D and ICT	0.028 (0.004)	67	52	0.023 (0.004)	0.021 (0.004)	3.50
Average salary in R&D and ICT	9.675 (1.406)	67	52	7.903 (1.349)	7.513 (1.390)	3.10
Average salary in wholesale-retail	9.506 (1.450)	67	52	3.795 (1.407)	3.014 (1.600)	1.30
Night lights (2009–2011)	1.547 (0.155)	78	59	0.533 (0.155)	0.431 (0.151)	1.70
Number of SMEs per 1,000 people	−3.106 (1.082)	65	50	0.243 (1.369)	0.757 (1.335)	1.00
SME labor productivity	0.813 (0.088)	65	50	0.492 (0.104)	0.439 (0.102)	2.15

Notes: Standard errors are reported in parentheses. For raw differences in the whole sample, the standard errors are the denominators of t-statistics of group differences. 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 a 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 $\alpha = 0.05$ type I error.

(see online Appendix Table D.1). Raw differences suggest that Science Cities are characterized by a lower SME density (number of SMEs divided by municipal population); the corresponding ATT estimates indicate the opposite but are only significant (at the 10 percent level) when adjusting for the matching bias. The SME labor productivity ATT estimate is, however, positive and statistically significant. This suggests that the economic effect of Science Cities operates on the intensive (productivity) margin of the local economy.

We also perform a sensitivity analysis of our ATT estimates. Following Rosenbaum (2002), we simulate unobserved factors that would affect both the outcomes and the probability that a municipality hosts a Science City. We assess to what extent this would influence our conclusions about the presence of statistically significant differences in Y_{iq} between treated and (matched) control observations, for all outcomes $q = 1, \ldots, Q$. The size of the simulated unobserved factor is given by parameter $\Gamma \geq 1$, which measures the hypothesized odds of receiving the treatment ($\Gamma = 1$ is the experimental benchmark). In Table 1 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 of the differences between treated and control observations.[25] The values of Γ^* are quite high (around 3 or more) for the patent applications, the employment share and average salary in R&D and ICT, and the graduate and postgraduate shares. They are satisfactorily high (1.7 or higher) for total population, night lights, and SME labor productivity.[26] These values are in line with our statistical inference on 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 and 1.3—for average salary in wholesale-retail and SME density: the conclusions about these outcomes appear less robust, although this could also be due to a reduced sample size thanks to ROSSTAT's incomplete coverage.

ATT Estimation: Historical Science Cities.—Our interpretation of these estimates in terms of long-run effects would be threatened if, on average, Science Cities still receive a preferential treatment from the Russian government, in the form of direct or indirect support to local R&D or general purpose expenditure, such as infrastructure. In order to assess to what extent our results depend on current governmental support, we repeat the above analysis, excluding Science Cities with the official status of Naukogrady in today's Russia. For these 14 Science Cities, the Russian government has resumed the Soviet-era program in the early 2000s, although with a less military and more civil focus. We call the remaining Science Cities "historical."

Table 2 shows the estimates based on the matched sample restricted to historical Science Cities, with the significance levels based on a FWER adjustment using Holm's method that groups all the outcomes listed in Table 2 together with the "demographic dynamics" outcomes listed in Table F.2 in online Appendix F (see Schweiger, Stepanov, and Zacchia 2022 for the stars associated with the standard levels of significance based on adjusted *p*-values). The estimated ATT is, for most outcomes of interest, very similar to the corresponding estimates in Table 1. Statistical inferences and sensitivity analyses a la Rosenbaum generally confirm this assessment. The removal of Naukogrady results in a dramatic change of the estimated effects for patenting activity and employment share in R&D and ICT. The measure relative to pre-2000 patent applications is about 20 percent smaller than the initial estimates in Table 1, while the measure relative to later patent applications is about 28 percent smaller. Employment share in R&D and ICT is about 1 percentage point lower than the initial estimate—this is substantial, considering that the raw mean employment share in R&D and ICT is 3.6 percent in Science Cities. Nevertheless, all estimates remain significant and robust, as evidenced by the values of Γ^* from the sensitivity analysis.

[25] The Wilcoxon signed-rank tests are based on the plain differences between matched pairs for every outcome; we select Γ^* as the smallest values of Γ in the grid such that any *p*-value of the test is higher than $\alpha = 0.05$. Full results of the sensitivity analysis are available on request.

[26] $\Gamma = 2$ 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.

Table 2—Municipal-Level Matching, Main Analysis: "Historical" Science Cities

	Whole sample	Matched sample (1 nearest neighbor)				
Outcome	Raw difference	*T*	*C*	ATT	ATT b.a.	Γ^*
Population in 2010 (thousands)	38.419 (10.740)	64	51	33.570 (9.458)	33.472 (7.696)	2.10
Graduate share in 2010	0.110 (0.009)	64	51	0.056 (0.008)	0.052 (0.009)	3.45
Postgraduate share in 2010	0.003 (0.000)	64	51	0.002 (0.000)	0.003 (0.000)	2.30
Fractional patents before 2000	2.031 (1.046)	64	51	2.369 (0.638)	2.229 (0.631)	2.85
Fractional patents since 2000	5.868 (2.551)	64	51	6.664 (1.874)	6.243 (1.865)	2.75
Employment share in R&D and ICT	0.018 (0.003)	53	44	0.012 (0.002)	0.011 (0.002)	2.45
Average salary in R&D and ICT	8.926 (1.271)	53	44	7.860 (1.144)	7.409 (1.353)	2.85
Average salary in wholesale-retail	8.853 (1.646)	53	44	3.649 (1.212)	3.420 (1.255)	1.15
Night lights (2009–2011)	1.313 (0.167)	64	51	0.494 (0.140)	0.428 (0.135)	1.50
Number of SMEs per 1,000 people (all)	−4.084 (1.117)	52	43	−1.381 (1.094)	−0.366 (1.147)	1.00
SME labor productivity (all)	0.751 (0.102)	52	43	0.453 (0.083)	0.388 (0.081)	2.20

Note: See the notes accompanying Table 1.

While the difference in the drop in patent applications before and after 2000 is not large, it suggests that current government support is an important stimulant for patenting activity in the Russian context, where innovation is still predominantly driven by the government sector.[27] This is underlined by the estimate for the employment share in R&D and ICT. In either case, we keep observing a positive differential in favor of historical Science Cities for most demographic and economic outcomes of interest. Such differentials are even more surprising, as they are independent of the extent to which the government currently supports local R&D and thus can only be interpreted as long-run effects. Our initial interpretation of the results is, if anything, reinforced by this restricted analysis.

ATT Estimation: Demographic Dynamics.—It is possible that our results are merely transient. In our conceptual framework (Section V), we postulate the existence of "persistence forces," independent of other endogenous mechanisms, that induce path-dependence from the Soviet-era allocation of the labor force. In the real world, however, workers are slowly replaced by new generations of younger

[27] A subsample analysis analogous to those reported later in Table 5 reveals that the difference-in-differences of post-2000 patents, calculated by comparing "historical" Science Cities against the official Naukogrady, is statistically significant at conventional levels; however, a similar test for pre-2000 patents is inconclusive. These results can be accessed through our online replication package.

Table 3—Municipal-Level Matching, Demographic Dynamics: All Science Cities

	Whole sample	Matched sample (1 nearest neighbor)				
Outcome	Raw difference	T	C	ATT	ATT b.a.	Γ^*
Graduate share: Birth year $\leq$ 1965	0.122	78	59	0.079	0.071	4.40
	(0.010)			(0.012)	(0.012)	
Graduate share: Birth year $>$ 1965	0.107	78	59	0.056	0.046	3.10
	(0.008)			(0.010)	(0.010)	
Postgraduate share: Birth year $\leq$ 1955	0.003	78	59	0.003	0.003	2.80
	(0.001)			(0.001)	(0.001)	
Postgraduate share: Birth year $>$ 1955	0.003	78	59	0.003	0.003	2.25
	(0.000)			(0.001)	(0.001)	

Note: See the notes accompanying Table 1.

workers. If the latter are increasingly less anchored to prior locations, spatial equilibrium can over time lead to mean reversion—even in the presence of agglomeration forces, thanks to the action of random shocks. 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 information on the number of residents in each municipality by type of attained education within each cohort of birth. This lets us assess the extent to which our results on urban educational levels are driven mainly by older cohorts. We split the population of each municipality observed in the 2010 census into the "young" and the "old," using two different threshold years of birth: 1965 for the graduate share and 1955 for the postgraduate share (taking into account the fact that in Russia, postgraduate education is characterized by a longer average duration). At the time of the dissolution of the USSR (1991–1992), the older individuals in the "young" group who had obtained a university degree were starting their professional careers and presumably could move more easily. Furthermore, those who were underage at the time of the transition might have pursued less education than their ancestors (mean reversion). Both factors would predict a more equal distribution of young graduates between Science Cities and their matched counterparts.

Using our matched sample, we estimate the Science Cities ATT on the graduate population share separately for the "old" and "young" groups. The results in Table 3 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 65 percent of the former.[28] The estimates are qualitatively similar for the postgraduate share of the population[29] and are not sensitive to the choice of the

[28] Significance is based on a FWER adjustment using Holm's method that groups all the outcomes listed in Table 3 together with the main outcomes listed in Table 1; see Schweiger, Stepanov, and Zacchia (2022) for the stars associated with the standard levels of significance based on adjusted p-values.

[29] 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. Across all municipalities, the unweighted average share of graduates in the old group is about 12.5 percent, compared with about 11 percent among the young (24.5 versus 21.5 percent in Science Cities). Conversely, the postgraduate share is 0.15 percent in the old group and 0.33 percent in the young group (0.50 versus 0.63 percent in Science Cities).

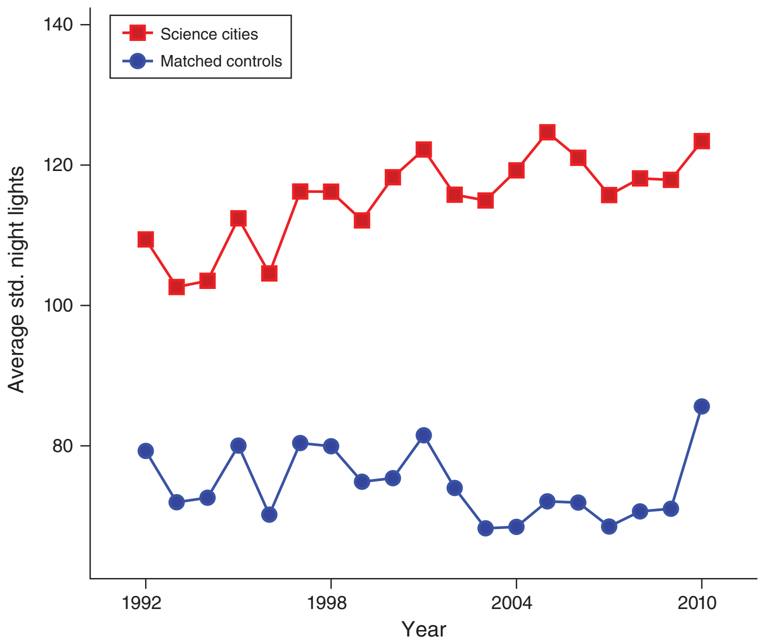

FIGURE 2. TIME SERIES OF THE AVERAGE STANDARDIZED NIGHT LIGHTS INDICATORS, 1992–2010

Source: Version 4 DMSP-OLS Nighttime Lights Time Series, National Oceanic and Atomspheric Administration (NOAA)

threshold. Thus, this analysis provides no evidence in favor of the mean reversion hypothesis: it appears that the children of Soviet inhabitants of Science Cities pursue educational and locational choices that are largely similar, albeit not identical, to those of their ancestors. Estimates for the restricted "historical" subsample (see online Appendix F) support these conclusions.

ATT Estimation: Economic Dynamics.—A logical next step is to assess mean reversion in economic outcomes. If the relative skill levels of Science Cities and comparable municipalities are equalized 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 posttransition years, with the exception of the night lights measures. Figure 2 depicts averages of the standardized night lights indicator, separately for Science Cities and their matched controls, for every year from 1992 to 2010. While the two groups share similar annual fluctuations, Science Cities appear to constantly outperform their counterparts, with hardly any catch-up by the control group. To check that this pattern is not due to an extreme path dependence of some random unobserved factors that are not explained by Science City status, we perform formal regression-based tests, allowing for temporal persistence in the unobservable factors driving each municipality's night light measures (see online Appendix F).

ATT Estimation: Municipal Budgets and Local Public Goods.—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 vES. They explain their results by the persistence of municipal spending in certain, presumably productivity-enhancing, infrastructure, rather than by agglomeration forces. 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 their resources from both local taxes (property taxes, merchant fees, fees for the provision of local services) and from a portion of federal taxes (income tax, business tax, and similar) that are paid by local residents. In addition, municipalities receive discretionary transfers from the federal and the regional governments. In our data, we can 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. Our measures of interest are 2006–2016 averages of selected budget items for each municipality, divided 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.[30]

Table 4 summarizes the key estimates. In raw differences, Science Cities collect more taxes per capita than ordinary municipalities; however, they receive disproportionately lower total transfers per capita. As a result, both their total revenues and expenditures per capita are smaller. This is only partly mitigated by the fact that Science Cities obtain higher earnings from local taxes. In the matched sample, the ATT estimates for average revenues and expenditures are equal to zero, those for average tax income are positive and statistically significant—as one would expect if Science Cities were indeed richer—and those for average transfers are negative and also statistically significant, with and without bias adjustment.[31] The values of Γ^* are equal to 1 for all outcomes except tax income, in which case it is close to 2.

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 Russia. Since Science Cities are typically richer and thus obtain higher local taxes, this often results in lower total transfers in their

[30] In this part, we control for the FWER within a family of variables that comprises both municipal budget and local public goods. We keep this family of "potential mechanisms" separate from the main outcome variables in Table 1 because they are meant to address different hypotheses (we could find differences in the potential mechanisms and not in the main outcomes, or vice versa). Inferences are not fundamentally altered if we group all outcomes in one family.

[31] Statistical significance is based on a FWER adjustment using Holm's method that groups all the outcomes listed in Table 4 together with the detailed budget entry outcomes and the "amenity" outcomes listed in Table F.1 in online Appendix F. See Schweiger, Stepanov, and Zacchia (2022) for the stars associated with the standard levels of significance based on adjusted *p*-values.

Table 4—Municipal-Level Matching, Budgets Analysis: All Science Cities

	Whole sample	Matched sample (1 nearest neighbor)				
Outcome	Raw difference	T	C	ATT	ATT b.a.	Γ^*
Total revenues per capita	−5.271	64	51	−0.283	−1.183	1.00
	(1.375)			(0.865)	(1.077)	
Tax income per capita	3.467	64	51	1.986	1.520	1.95
	(0.726)			(0.473)	(0.527)	
All transfers per capita	−8.738	64	51	−2.269	−2.703	1.00
	(0.869)			(0.688)	(0.850)	
Total expenditures per capita	−5.158	64	51	0.236	−0.340	1.00
	(1.357)			(0.871)	(1.063)	

Notes: See the notes accompanying Table 1.

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

To test the vES hypothesis in more detail, 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 find no evidence of differential patterns in any of the budget subcategories we examine; none of the ATT estimates are statistically significant, and the corresponding values of Γ^* are equal to 1. In addition, we test if Science Cities differ from their matches in terms of the existing stock of selected public goods, like libraries, museums, theaters (facilities that are more likely to attract the better educated) or the shares of streets lit during the night. Once again, we find no robust differences. The detailed results can be found in online Appendix F, which also reports the results from the analysis of municipal budgets restricted to the subsample of "historical" Science Cities—an exercise that leaves our main conclusions unchanged.

Robustness: Effects by Science City Categories.—Lastly, we explore whether our results are driven by specific features of Science Cities or of the matched controls. In particular, we consider the following categorizations. Reminiscent of the security-usability trade-off discussed in Section I, we first inquire whether Science Cities originally serving top secret military purposes—those in aviation/rocket or in nuclear science—have embarked on a different pattern of economic development thanks to, say, more robust past investments. Second, we look at Science Cities that were built from scratch vis-à-vis those created in pre-existing settlements; the pattern of demographic and economic development might have been different in the two groups.[32] Third, we examine whether the long-run outcomes of the Science Cities

[32] For example, Science Cities built from scratch may have initially grown faster but have stagnated following the transition in case they had been placed in inconvenient locations. Note that Science Cities built from scratch

Table 5—Robustness t-Tests: Differences between Types of Science Cities

Criterion	Fractional patents since 2000	Night lights	Total population	Graduate share
Secrecy versus usability	−0.007	1.861	2.235	1.445
	[0.995]	[0.136]	[0.125]	[0.205]
Built from scratch versus others	0.178	−0.560	−0.520	−0.654
	[0.859]	[0.808]	[0.808]	[0.808]
Agirrechu's list versus others	−1.267	0.636	1.980	−0.075
	[0.425]	[0.706]	[0.213]	[0.941]
Other close Science Cities, versus none	1.261	−0.715	−0.697	−0.044
	[0.651]	[0.651]	[0.651]	[0.965]
Founded before 1950, versus later	1.168	0.103	1.378	0.152
	[0.493]	[0.918]	[0.493]	[0.918]
Founded before 1960, versus later	1.003	−0.182	0.105	−0.028
	[0.978]	[0.978]	[0.978]	[0.978]

Notes: For every combination given by a Science City category (rows) and an outcome variable (columns), the table displays the t-statistic of the average treated-control mean differences in the outcome between the two parts of the matched sample, and (in brackets) the corresponding adjusted p-value that controls for the FDR across all four outcomes for each criterion, using Hochberg's method. Online Appendix G provides a visual representation of the differences being tested.

identified by Agirrechu (2009), whose list we consider incomplete and best amended through other sources, are in any way different from those of the remaining cities in our list.[33] Fourth, we provide a test about possible violation of the SUTVA by splitting the matched sample according to the presence of other Science Cities within a radius of 50 km from the matched control, or lack thereof.[34] Finally, the creation of Science Cities was staggered in the postwar decades, but we match mostly using covariates measured in 1959. This could potentially lead us to understate the effect on early Science Cities and overstate the effect on the later ones. To analyze whether the timing of their creation has any bearing, we use 1950 or 1960 as cutoff dates.

To provide an assessment about differences between the cited categories, we perform t-tests of a difference in the treated-control means in the matched sample. Given the small size of some of the subsamples in question, we restrict our attention to key outcomes for which we have complete information for all municipalities: the post-2000 fractional patent applications, the night lights measure averaged over 2009–2011, the 2010 population, and the share of that population with at least a graduate degree.

Table 5 summarizes the results of these t-tests, for each category-outcome combination, showing the t-statistic value and adjusted p-value controlling for the false discovery rate (FDR) for the outcomes in each category, using Hochberg's method.[35] None of the null hypotheses can be rejected at conventional levels when controlling for

were newly urbanized settlements located in a municipality that included other, possibly sparsely populated, settlements; under our matching approach, these Science Cities are paired to municipalities with similar characteristics and pre-trends.

[33] The original list by Agirrechu (2009) includes 75 Science Cities out of our total of 95.

[34] Our matched sample is constructed by imposing a constraint of a 50 km minimal distance between treated and control cities. Therefore, under the hypothesis that interference only affects control cities and it does not affect places farther than 50 km, SUTVA violation is not a concern for matched pairs where no other Science Cities are too close to the control municipality (39 out of 78 in our matched sample). See online Appendix F for additional discussion.

[35] We control for the FDR rather than FWER to make our tests more powerful at detecting true positives.

the FDR. Overall, we conclude that the specific categories of Science Cities we examined and issues of SUTVA violation are unlikely to drive our main results.

V. Model and Mechanisms

To enhance the interpretation of the long-run effect of Science Cities, we estimate a spatial equilibrium model suited to our particular setting. We adapt a simplified version of the framework by Allen and Donaldson (2020), while introducing workers of two different skill types. We also draw inspiration from Moretti (2011, 2013); Glaeser and Gottlieb (2008, 2009); and prior tradition (Rosen 1979; Roback 1982). We first describe the model and its equilibrium properties and then discuss the resulting estimates and their implications. A detailed analysis of the model and its extensions can be found in online Appendix H.

Model Setup.—Consider a set $\mathcal{C}$ of Russian cities, of dimension C and indexed as $c = 1, \ldots, C$. These cities are inhabited by two different types of workers: those of high skill and those of relatively lower skill. This classification is conventionally interpreted in terms of differences in the educational attainment. In our context, highly skilled workers can be identified as researchers engaged in R&D—a subset of all the university-educated—while low-skilled workers represent all other workers in the remaining sectors. The model allows for both interpretations. We denote the mass of highly skilled workers employed in city c at time t as H_{ct}, while L_{ct} is the corresponding notation for low-skilled workers. We use lowercase letters, respectively h_{ct} and ℓ_{ct}, to denote their logarithms.

At time $t = 0$, all cities are part of the Soviet Union, which, for idiosyncratic reasons, allocates the values of H_{ct} and L_{ct} to each city. At time $t = 1$, all cities are part of modern Russia, a market economy, and workers of both types self-select into either location. We express the logarithmic indirect utility u_{nic} of an individual i of type $n \in h, \ell$ living in city c as a linear function of wages, local characteristics such as amenities, past and present population of either skill type, and idiosyncratic preferences:

$$u_{ic}^{n} = w_{c}^{n} + \beta_{0}^{n}\ell_{c0} + \beta_{1}^{n}\ell_{c1} + \delta_{0}^{n}h_{c0} + \delta_{1}^{n}h_{c1} + a_{c}^{n} + e_{ic}^{n}, \tag{1}$$

where w_c^n is the log-wage earned by workers of type n in city c at $t = 1$; a_c^n represents factors specific to city c (such as amenities or public goods) that make it a more enjoyable location, and whose specific value might vary by skill group; and e_{ic}^n denotes the idiosyncratic taste of individual i for city c—a random shock assumed to follow a Type I Extreme Value (Gumbel) distribution with zero location parameter and scale parameter equal to σ^n. In this expression, parameters β_1 and δ_1 represent current population externalities, which could be negative if "divergence" effects dominate. Instead, β_0 and δ_0 express the strength of those "persistence" forces that make a location desirable by virtue of its past population.[36] If $\sigma^h - \delta_1^h > 0$ and

[36] Typical examples of divergence forces are the "congestion" effects due to traffic or pollution and higher rents due to a higher demand for housing. Persistence forces can be micro-founded on better housing stocks and/or wider product variety that are due to a higher past population.

$\sigma^\ell - \beta_1^\ell > 0$, the model generates positive labor supply elasticities at the city pair level.

To close the model, we introduce two types of firm: one that employs highly skilled workers and one that relies on lower-skilled workers.[37] The log-output y_c^n of type $n = h, \ell$ firms in city c follows a Cobb-Douglas technology:

$$y_c^h = g_c^h + \theta^h h_c + \varphi^h \ell_c + \lambda h_c + (1-\lambda) k_c^h \tag{2}$$

$$y_c^\ell = g_c^\ell + \theta^\ell h_c + \varphi^\ell \ell_c + \lambda \ell_c + (1-\lambda) k_c^\ell,$$

where g_c^n is the city- and type-specific total factor productivity and k_c^n 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.[38] For simplicity, the elasticity of labor is equal to $\lambda \in (0,1)$ for both types of firm in both cities. Firms of type ℓ do not hire workers of type h but take h_c as given; vice versa, firms of type h take ℓ_c as given. The parameters θ^h and θ^ℓ represent agglomeration forces, such as knowledge spillovers, that stem from a higher high-skilled population and affect firms employing high- and low-skilled workers, respectively. Symmetrically, φ^h and φ^ℓ represent agglomeration forces due to a higher low-skilled population.[39] This specification nests a more restricted one where $\vartheta^n \equiv \theta^n = -\varphi^n$ for $n = h, \ell$, in which case, agglomeration effects are a function of the share of high-skilled workers in city c, similar to Moretti (2004).

Spatial Equilibrium.—The spatial equilibrium is most conveniently expressed as the log-difference in selected outcomes between any two cities s and z in $\mathcal{C}$. The key equilibrium equations are, for $n = h, \ell$, as follows:

$$(n_{s1} - n_{z1}) = (p_s^n - p_z^n) + \mu^{nh}(h_{s0} - h_{z0}) + \mu^{n\ell}(\ell_{s0} - \ell_{z0}) \tag{3}$$

$$(w_s^n - w_z^n) = (r_s^n - r_z^n) + \frac{\theta^n \mu^{hh} + \varphi^n \mu^{\ell h}}{\lambda}(h_{s0} - h_{z0}) + \frac{\theta^n \mu^{h\ell} + \varphi^n \mu^{\ell\ell}}{\lambda}(\ell_{s0} - \ell_{z0}), \tag{4}$$

where p_c^n and r_c^n are functions of g_c^n and a_c^n for $n = h, \ell$, while the parameters expressed as $\mu^{nn'}$—where $n' = h, \ell$—are functions of all the parameters listed in (1) and (2), including σ^h, σ^ℓ. We call these parameters "history multipliers": they express the equilibrium elasticities of present $(t = 1)$ population of type n on the past $(t = 0)$ population of type n'. As such, they enclose all the interaction effects

[37] In Moretti (2011), this was largely a simplification meant to abstract from the degree of substitutability between skills. This characteristic of the model can be given a contextual interpretation here: if workers of type h are researchers, type h firms correspond to the R&D sector, while type ℓ firms represent the rest of the local economy.

[38] See the discussion in online Appendix H about the realism of this hypothesis and the implications of its violation.

[39] While Allen and Donaldson (2020) do not distinguish agglomeration forces due to different skill types, they include agglomeration forces that depend upon *past* population (although they find these to be very small in their empirical analysis). We omit them, as we are unable to identify the implied additional parameters with the data available to us.

of agglomeration and divergence forces that lead to the spatial dispersion of population and economic activity. As detailed in the online Appendix, to achieve a nondegenerate equilibrium, it is necessary that none of these forces is too prevalent, implying some restrictions on the values of the primitive parameters.

Estimation of the Model.—We augment p_c^n and r_c^n by making them both functions of observable city characteristics (like local expenditures in public goods, or amenities) and error terms. Hence, the four equations given by (3) and (4) for $n = h, \ell$ are reshaped as the reduced forms of two pairs of triangular equations with two endogenous right-hand-side variables (the differences in $t = 1$ populations for both skill types) and two exogenous ones (the differences in $t = 0$). We estimate the model on our matched sample so that s denotes a Science City and z its matched counterpart. This aids identification: under our maintained hypotheses, within the matched sample the Soviet-era allocation of h_c and ℓ_c for $c = s, z$ is orthogonal to secular determinants of socioeconomic outcomes. Note that the municipal-level population data from 1989 that we collected cannot be distinguished by level of education. To obviate this problem, we adopt two complementary approaches. In the first, we define "past" and "present" populations as the individuals from the 2010 census who were born before and since 1965, respectively (Approach 1). In the second, we proxy the two exogenous variables by the fitted values obtained from regressing 1989 log-population on the "old" 2010 log-population of either skill type (Approach 2). The assumption implicit in both approaches is that older individuals are immobile. Lastly, since λ is not identified in our model, we calibrate it as $\lambda = 0.66$.

The results are displayed in Table 6. The two approaches deliver very similar estimates. In the baseline estimates in columns 1 and 3, the high-to-high spillover elasticity θ^h is estimated at around 0.27 and is significant at the 1 percent level; the estimates for the other parameters are less robust. The baseline estimates cannot reject the restriction that $\vartheta^n \equiv \theta^n = -\varphi^n$ for $n = h, \ell$; consequently, in columns 2 and 4, we also show estimates that incorporate this restriction. Under both approaches, the estimates of parameter ϑ^h—the elasticity of high-skilled wages on the share of high-skilled workers in a city—are around 0.33 and statistically significant. The corresponding estimates ϑ^ℓ for the low-skilled are smaller in magnitude and less precisely estimated. The estimates of the history multipliers are significant, with the exception of the high-on-low elasticity $\mu^{h\ell}$. This suggests that while the sizes of both skill groups depend on their past values, a higher high-skilled population attracts more low-skilled workers in the future, but the converse is not true.

Extrapolations.—The estimates from the model help to enhance the interpretation of our main results from Section IV in two major ways. First, they allow us to evaluate the relative effects of alternative counterfactual allocations of labor from $t = 0$ on two pairs of matched cities s and z. In particular, given an allocation $(h_{s0}, h_{z0}, \ell_{s0}, \ell_{z0})$, one can calculate the following quantity, for $n = h, \ell$:

$$\mathcal{W}_{sz}^n = \exp\big(\big(w_s^n + n_{s1}\big) - \big(w_z^n + n_{z1}\big)\big), \tag{5}$$

Table 6—Estimation of the Spatial Equilibrium Model

	Approach 1		Approach 2	
Parameter	(1)	(2)	(3)	(4)
$\theta^h(\vartheta^h)$: high-to-high spillovers	0.269	0.335	0.249	0.323
	(0.061)	(0.059)	(0.053)	(0.063)
φ^h: low-to-high spillovers	−0.211		−0.191	
	(0.087)		(0.079)	
$\theta^\ell(\vartheta^\ell)$: high-to-low spillovers	0.155	0.175	0.142	0.165
	(0.077)	(0.089)	(0.070)	(0.092)
φ^ℓ: low-to-low spillovers	−0.136		−0.124	
	(0.104)		(0.096)	
μ^{hh}: history: high-on-high	0.976	0.942	1.683	1.609
	(0.031)	(0.044)	(0.029)	(0.074)
$\mu^{h\ell}$: history: high-on-low	0.056	0.123	0.031	0.127
	(0.051)	(0.079)	(0.036)	(0.094)
$\mu^{\ell h}$: history: low-on-high	0.110	0.145	0.108	0.181
	(0.018)	(0.042)	(0.020)	(0.075)
$\mu^{\ell\ell}$: history: low-on-low	0.902	0.835	1.212	1.116
	(0.025)	(0.068)	(0.021)	(0.090)
Spillover constraints test: p-value	0.378		0.385	
Full set of covariates	Yes	Yes	Yes	Yes
Constrained model	No	Yes	No	Yes
Observations	55	55	55	55

Notes: The table displays estimates of the model expressed by equations (3) and (4) for $n = h, \ell$, where g_c^n and a_c^n are expressed as linear functions of a set of city characteristics ("covariates") and an error term as discussed in online Appendix H, and for both approaches described in the text. Estimation is performed on the matched sample via one-step GMM using an identity weighting matrix. In constrained models, the parameter restrictions $\vartheta^n \equiv \theta^n = -\varphi^n$ for $n = h, \ell$ are applied; in unconstrained models, the p-value from a joint test of these restrictions ("spillover constraints test") is reported. Standard errors are clustered by grouping matched pairs that share the same control city; they are reported in parentheses.

which represents the ratio between the total wages of type n paid in city s and those paid in city z. To substantiate, we provide values of $\mathcal{W}_{sz}^n$ for both skill types, calculated using the estimates from column 4 of Table 6, for six alternative allocations. The results from this exercise are displayed in Table 7.

Column 1 represents a baseline scenario where Science Cities had never been established: $H_{c0} = 14$ and $N_{c0} = 68$ (in thousands of inhabitants) for $c = s, z$. Column 2 corresponds to a representative allocation of high-skilled workers from other, nonmatched locations to Science Cities, with H_{sz} increasing to 22 relative to column 1.[40] A way to interpret the predicted values of $\mathcal{W}_{sz}^n$ is by taking "ratios of ratios": for example, the total wages paid to high-skilled workers in a typical Science City (column 2) are about 60 percent higher relative to the nonintervention scenario (column 1), which is close to the actual difference that we observe in the matched

[40] We choose these numbers because they are roughly consistent with the average values of the 1989 population in Russia (90 in Science Cities and 82 in their matched controls), with a graduate share among old generations of about 0.24 in Science Cities and of 0.16 in their matched controls (as seen in the data) and with the actually observed values of $\mathcal{W}_{sz}^h$ and $\mathcal{W}_{sz}^\ell$ in the 2010 data.

Table 7—Wage Ratios Defined in Equation (5) for Alternative Allocations at $t = 0$

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathcal{W}_{sz}^{h}$	1.000	1.606	1.301	2.223	0.974	1.442
$\mathcal{W}_{sz}^{\ell}$	1.000	1.183	1.098	1.328	1.112	1.827
(H_{s0}, L_{s0})	(14,68)	(22,68)	(18,68)	(30,68)	(14,76)	(22,88)
(H_{z0}, L_{z0})	(14,68)	(14,68)	(14,68)	(14,68)	(14,68)	(14,56)

Notes: In each column, the table provides the values of $\mathcal{W}_{sz}^{h}$ and $\mathcal{W}_{sz}^{\ell}$ that are predicted by the estimates from column 4 of Table 6 given the allocation $(H_{s0}, H_{z0}, L_{s0}, L_{z0})$, expressed in thousands of inhabitants, specified at the bottom. See online Appendix H for the formulas used to produce these values.

sample. One can perform similar exercises using other figures from Table 7.[41] Note that these calculations cannot reveal the loss that other Russian cities might have incurred relative to the counterfactual. To provide a comprehensive answer about the overall welfare effects of such a policy, it is necessary to develop and estimate a more extensive model, such as Rossi-Hansberg, Sarte, and Schwartzman (2021). They study a hypothetical policy intervention in the United States that focuses on concentrating cognitive nonroutine occupations in cognitive hubs and argue that it would improve welfare by optimally harnessing the externalities. Assessing the welfare effects of an actually tried policy that is similar in spirit, such as the Science Cities evaluated in this paper, would be a good direction for future work.

Second, the parameters of the model can be used to calculate the subsidy t_N^* (expressed in units of log-wage) that the government should provide as an incentive to workers of type $N = H, L$ for them to move to some city s so as to reproduce ex post the population ratio $N^* = N_s/N_z$, where s and z are two ex ante identical cities. This value is lower than one implied by the city-level labor supply elasticity because the effect of a relatively small incentive is amplified in equilibrium through agglomeration economies. However, the parameter estimates from Table 6 are not enough to perform this calculation: it is also necessary to know the parameters β_t^n, δ_t^n, and σ^n, for $t = 0, 1$ and $n = h, \ell$. In online Appendix H, we discuss an attempt to use external information—drawn from the literature—to calculate the subsidy t_H^* that the Russian government should provide to reproduce the relative allocation typical of Soviet Science Cities: $(h_{0s} - h_{0z}) = \log(22/14)$, in today's market economy. In summary, we calculate that the government should provide a subsidy equal to about 150 percent of the ex ante wage; this figure is based on the city-level labor supply elasticities of about 0.26. In future work we plan to obtain elasticity estimates based on cross-municipality migration data; this would allow us to provide more accurate values for the subsidy t_N^*.

[41] For example, column 3 predicts the consequences of an intervention of half the size relative to the baseline; column 4, of double size; column 5 describes an alternative intervention of increasing the allocation of low-skilled workers, rather than high-skilled ones, to city s; and finally column 6 describes an allocation where the s-z skill type ratios are identical.

VI. Conclusion

In this paper, we have analyzed the long-run effects of a unique historical place-based policy: the creation of R&D-focused Science Cities in Soviet Russia. Both the initial establishment and the eventual suspension of this program was largely guided by political factors that are arguably exogenous to drivers of current social and economic conditions of Russian cities. We compare Science Cities with other localities that were observationally similar to them at the time of their selection, and we compute differences in the current characteristics between the two groups. We find that former Science Cities are bigger today and they host a higher number of well-educated individuals. Moreover, they file a higher number of patent applications at the EPO; their R&D and ICT sectors are more developed and pay higher salaries. Lastly, Science Cities host more productive small businesses.

Because our results are largely unchanged after we exclude Science Cities that currently receive support from the Russian government, we conjecture that they are driven by the interaction between persistence (path-dependence) and agglomeration forces. In more mundane terms, highly skilled individuals who have remained in their former cities of residence have contributed to the emergence of more productive businesses in the new market economy. This insight is substantiated by the estimates of agglomeration elasticities based on a spatial equilibrium model. We rule out alternative explanations that have to do with differential governmental transfers and find no support for the hypothesis of rapid mean reversion of socioeconomic outcomes.

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

REFERENCES

Abadie, Alberto, and Guido W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74 (1): 235–67.

Abadie, Alberto, and Guido W. Imbens. 2011. "Bias-Corrected Matching Estimators for Average Treatment Effects." *Journal of Business and Economic Statistics* 29 (1): 1–11.

Agirrechu, Aleksandr A. 2009. *Russian Science Cities: History of Formation and Development* (*Naukogradi Rossiyi: Istoriya formirovaniya i razvitiya*). Moscow: Moscow University Press.

Allen, Treb, and Dave Donaldson. 2020. "Persistence and Path Dependence in the Spatial Economy." NBER Working Paper 28059.

Ambler, John, Denis J.B. Shaw, and Leslie Symons, ed. 1985. *Soviet and East European Transport Problems*. New York: St. Martin's Press.

Audretsch, David B., and Maryann P. Feldman. 1996. "R&D Spillovers and the Geography of Innovation and Production." *American Economic Review* 86 (3): 630–40.

Ben-Michael, Eli, Avi Feller, and Jesse Rothstein. 2020. "The Augmented Synthetic Control Method." Unpublished.

Bircan, Çağatay, and Ralph De Haas. 2020. "The Limits of Lending? Banks and Technology Adoption across Russia." *Review of Financial Studies* 33 (2): 536–609.

Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica* 81 (4): 1347–93.

Cheremukhin, Anton, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinski. 2017. "The Industrialization and Economic Development of Russia through the Lens of a Neoclassical Growth Model." *Review of Economic Studies* 84 (2): 613–49.

Cooper, Julian M. 2012. "Science-Technology Policy and Innovation in the USSR [http://www.ceelbas.ac.uk/workshops/skolkovo/Cooper]." CEELBAS Workshop "Russia's Skolkovo in Comparative and Historical Perspective", held on 12 June 2012 at UCL SSEES. Slides.

Diamond, Alexis, and Jasjeet S. Sekhon. 2013. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." *Review of Economics and Statistics* 95 (3): 932–45.

Donaldson, Dave, and Adam Storeygard. 2016. "The View from Above: Applications of Satellite Data in Economics." *Journal of Economic Perspectives* 30 (4): 171–98.

Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. 2010. "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns." *American Economic Review* 100 (3): 1195–1213.

Fan, Jingting, and Ben Zou. 2021. "Industrialization from Scratch: The 'Construction of Third Front' and Local Economic Development in China's Hinterland." Unpublished.

Ganguli, Ina. 2015. "Immigration and Ideas: What did Russian Scientists 'Bring' to the United States?" *Journal of Labor Economics* 33 (S1): S257–88.

Glaeser, Edward L., and Joshua D. Gottlieb. 2008. "The Economics of Place-Making Policies." *Brookings Papers on Economic Activity* 39 (1): 155–239.

Glaeser, Edward L., and Joshua D. Gottlieb. 2009. "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States." *Journal of Economic Literature* 47 (4): 983–1028.

Glaeser, Edward L., Hedi D. Kallal, José A. Scheinkman, and Andrei Schleifer. 1992. "Growth in Cities." *Journal of Political Economy* 100 (6): 1126–52.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti. 2010. "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings." *Journal of Political Economy* 118 (3): 536–98.

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

Gregory, Paul, and Mark Harrison. 2005. "Allocation Under Dictatorship: Research in Stalin's Archives." *Journal of Economic Literature* 43 (3): 721–61.

Gregory, Paul R., and Andrei Markevich. 2002. "Creating Soviet Industry: The House That Stalin Built." *Slavic Review* 61 (4): 787–814.

Harrison, Mark. 2017. "Secrecy and State Capacity: A Look Behind the Iron Curtain." Warwick Economics Research Papers 1134.

Heblich, Stephan, Marlon Seror, Hao Xu, and Yanos Zylberberg. 2020. "Industrial Clusters in the Long Run: Evidence from Million-Rouble Plants in China." Unpublished.

Ivanov, Denis. 2016. "Human Capital and Knowledge-Intensive Industries Location: Evidence from Soviet Legacy in Russia." *Journal of Economic History* 76 (3): 736–68.

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

Kline, Patrick, and Enrico Moretti. 2014a. "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." *Quarterly Journal of Economics* 129 (1): 275–332.

Kline, Patrick, and Enrico Moretti. 2014b. "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs." *Annual Review of Economics* 6: 629–62.

Marshall, Alfred. 1920. *Principles of Economics*. 8th ed. London: Macmillan.

Mikhailova, Tatiana. 2012a. "Gulag, WWII and the Long-run Patterns of Soviet City Growth." Munich Personal RePEc Archive Paper 41758.

Mikhailova, Tatiana. 2012b. "Where Russians Should Live: A Counterfactual Alternative to Soviet Location Policy." Munich Personal RePEc Archive Paper 35938.

Moretti, Enrico. 2004. "Workers' Education, Spillovers and Productivity: Evidence from Plant-Level Production Functions." *American Economic Review* 94 (3): 656–90.

Moretti, Enrico. 2011. "Local Labor Markets." In *Handbook of Labor Economics*, Vol. 4B, edited by Orley Ashenfelter and David Card, 1237–1314. Amsterdam: Elsevier.

Moretti, Enrico. 2013. "Real Wage Inequality." *American Economic Journal: Applied Economics* 5 (1): 65–103.

Moretti, Enrico, Claudia Steinwender, and John Van Reenen. 2021. "The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers." NBER Working Paper 26483.

Neumark, David, and Helen Simpson. 2015. "Place-Based Policies." In *Handbook of Regional and Urban Economics*, Vol. 5A, edited by Gilles Duranton, J. Vernon Henderson, and William Strange, 1197–1288. Amsterdam: Elsevier.

Roback, Jennifer. 1982. "Wages, Rents and the Quality of Life." *Journal of Political Economy* 90 (6): 1257–78.

Rosen, Sherwin. 1979. "Wage-Based Indexes of Urban Quality of Life." In *Current Issues in Urban Economics*, edited by Peter N. Miezkowski and Mahlon R. Straszheim, 74–104. Baltimore: Johns Hopkins University Press.

Rosenbaum, Paul R. 2002. *Observational Studies*. 2nd ed. New York: Springer.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Felipe Schwartzman. 2021. "Cognitive Hubs and Spatial Redistribution." Unpublished.

Rowland, Richard H. 1996. "Russia's Secret Cities." *Post-Soviet Geography and Economics* 37 (7): 426–62.

Saltykov, Boris G. 1997. "The Reform of Russian Science." *Nature* 388: 16–18.

Schweiger, Helena, Alexander Stepanov, and Paolo Zacchia. 2022. "Replication data for: The Long-Run Effects of R&D Place-Based Policies: Evidence from Russian Science Cities." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E140861V1.

Siddiqi, Asif A. 2008. "Soviet Space Power During the Cold War." In *Harnessing the Heavens: National Defense Through Space*, edited by Paul G. Gillespie and Grant T. Weller, 134–50. Chicago: Imprint Publications.

Siddiqi, Asif. 2015. "Scientists and Specialists in the Gulag: Life and Death in Stalin's Sharashka." *Kritika: Explorations in Russian and Eurasian History* 16 (3): 557–88.

Toews, Gerhard, and Pierre-Louis Vézina. 2020. "Enemies of the People." Quantitative Political Economy Research Group Working Paper 2020–20.

Von Ehrlich, Maximilian, and Tobias Seidel. 2018. "The Persistent Effects of Place-Based Policy: Evidence from the West-German Zonenrandgebiet." *American Economic Journal: Economic Policy* 10 (4): 344–74.