The Geography of Inventors and Local Knowledge Spillovers in R&D*

Brian C. Fujiy†

U.S. Census Bureau

[Click here for the latest version of the paper]

Abstract

I causally estimate spatial knowledge spillovers in Research and Development (R&D) and quantify their importance for R&D policies. Using a new administrative panel on German inventors, I estimate these spillovers by isolating quasi-exogenous variation from the arrival of East German inventors across West Germany after the Reunification of Germany in 1990. I find that increasing the number of inventors working in a technological cluster by 10% leads to average inventor productivity gains of 4.09%. After embedding these spillovers into a spatial model of innovation, the model predicts that reducing inventor migration costs increases aggregate output significantly, and that the effectiveness of this policy increases with the degree of spillovers. The model also predicts that R&D subsidies should be place-based, and increasing with location fundamentals for R&D.

Keywords: Inventors, Research and Development, Innovation, Agglomeration, Spillovers

JEL Codes: D21, F16, J61, O31, O4

*This paper is based on the first chapter of my Ph.D. dissertation at the University of Michigan. I am grateful to my committee Andrei Levchenko, Jagadeesh Sivadasan, Sebastian Sotelo, and Dmitry Stolyarov for their guidance and support. Thanks for the comments and insights from Jose Asturias, Dominick Bartelme, Jaedo Choi, Nadim Elayan, Gaurav Khanna, Jose Moran, Dev Patel, Heitor Pellegrina, and Linda Tesar, as well as participants at Midwest Macro 2024, Harvard Business School, U.S. Census Bureau, and the University of Michigan. Thanks to Nikola Sander for sharing data on internal migration flows. Thanks to the staff of the Research Data Centre of the German Federal Employment Agency at the Institute for Employment Research (IAB-FDZ), and to Arun Mathur for providing me access to the data at UM. Any opinions and conclusions expressed herein are those of the author and do not represent the views of the U.S. Census Bureau.

†Email: bfujiy@gmail.com | Website: `https://bcfujiy.github.io/`

Introduction

Research and Development (R&D) is crucial for aggregate productivity due to its direct impact on innovation. At the same time, R&D exhibits substantially higher levels of spatial concentration than overall economic activity. For example, in 2014 in West Germany, a worldwide innovation powerhouse (WEF, 2018), around 30% of mechanical engineers worked in the top three cities in this profession. In comparison, only around 18% of workers located in the three most populated cities. Since Marshall (1890), agglomeration economies—spatial and inter-temporal knowledge spillovers, labor pooling, and customer-supplier linkages—are the core explanation for why economic activity concentrates. Nevertheless, the extent, causes and consequences of *local knowledge spillovers in R&D*—local productivity gains from the agglomeration of R&D activity—remain elusive. In this paper, I address the following research questions: (i) is there evidence of local knowledge spillovers in R&D? and (ii) are they quantitatively important for aggregate productivity when implementing R&D policies?

Addressing these questions is crucial to implement policies that promote economic activity through R&D. Governments around the world implement a variety of policies—reducing mobility or transportation costs, formation of economic clusters, among others—that leverage knowledge spillovers for their effectiveness (Feldman and Kelley, 2006). In particular, policies that promote R&D rely strongly on the local knowledge spillovers in this sector (Trajtenberg, 2001). Moreover, implementing these policies can generate general equilibrium effects due to the internal mobility of agents. Therefore, the design of policies that promote R&D activities requires both well-identified estimates of local knowledge spillovers in R&D, and a quantitative framework that accounts for these spillovers in general equilibrium. In this paper, I provide such estimates and framework, and apply them to study policies that promote R&D activities in Germany.

In this paper, I show that local knowledge spillovers in R&D are large and important for aggregate productivity. First, using new data on German inventors, I causally estimate such spillovers by isolating quasi-exogenous variation from the arrival of East German inventors across West Germany after the Reunification of Germany in 1990. I find that a 10% increase in the number of inventors in a location leads to average inventor productivity gains of around 4.09%. Second, I build a quantitative spatial model of innovation that account for the local knowledge spillovers in R&D I estimated in the data. Third, I calibrate the model and use it to quantify the productivity gains from implementing policies that promote R&D activities. I find that a 25% reduction of migration costs for inventors increases aggregate productivity by 5.87%, and that the 25% subsidy for firms' expenditures in R&D within the 2020 German R&D Tax Allowance Act would increase aggregate productivity by 4.27%. Finally, the productivity gains from these policies increase with the level of local knowledge spillovers in R&D. I now describe each of these steps in detail.

In the first part of the paper, I estimate the additional productivity that inventors gain from agglomerating. To perform this task, I leverage a matched administrative data on German inventors between 1980 and 2014. This data exhibits two features that makes it suitable for this paper. First, the dataset includes all the patents and their characteristics that inventors filed over time, so I can calculate the total number of forward citations of inventor's filed patents during a given period—inventor productivity—Second, the dataset tracks how inventors move across locations over time, so I can calculate cluster size as the number of inventors working in a given technological cluster, where a cluster is a technological area-location pair. An example of a cluster is mechanical engineering in Munich.

Then, I leverage variation in cluster size and inventor productivity to estimate the local knowledge spillovers in R&D; that is, whether a higher concentration of inventors leads to more productive inventors due to local knowledge spillovers.[1] The analysis compares inventors that moved to clusters of different sizes, and inventors that did not move but the number of inventors in the cluster changed. After saturating the model with a large set of fixed effects, I find that a 10% increase in cluster size is associated with average inventor productivity gains of around 1.75%. This elasticity is statistically significant at the 1%,

[1] Examples of how these spillovers manifest in the real world are interactions and exchange of ideas between inventors (Davis and Dingel, 2019).

and its significance is robust to different specifications of inventor productivity and time aggregation.

I then address potential endogeneity concerns that potentially biases when estimating local knowledge spillovers in R&D. For example, unobserved inventor idiosyncratic shocks could induce upward or downward biases. For example, if novice inventors systematically sort into large clusters, then spillovers estimates would suffer from downward biases. Also, unobserved cluster shocks could induce upward biases. For example, growth expectations in a technological cluster could increase both the productivity of inventors in that cluster and pull inventors into the cluster, and therefore induce an upward bias. Finally, measurement error could also introduce a downward bias.

To address these concerns, I propose an instrumental variable based on the historical episode of the Reunification of Germany in 1990. In particular, I leverage this natural experiment to construct a shift-share instrument that induces quasi-exogenous variation in the size of West German clusters, which I then use to causally estimate local knowledge spillovers in R&D. The "shifts" are leave-out shocks that measure the total number of inventors that moved from each location in East Germany towards any West German cluster, except to the instrumented cluster. The identification assumption is that these shocks are as-good-as-randomly assigned (Borusyak et al., 2022); that is, the shifts are uncorrelated with unobservables within the instrumented cluster. The strategy of leaving out the instrumented cluster from the construction of the shifts ensures that the shocks are constructed based solely on push factors arising from the East, and not from pull factors coming from the instrumented cluster. These shocks are then weighted by exposure "shares" that help predicting the number of inventors that move from each East German location to each West German cluster. These shares are constructed based on the inverse of the geographic distance between every location between East and West Germany, and the specialization of each location in East Germany in each technological area. Under this approach, a 10% increase in cluster size leads to average inventor productivity gains of around 4.09%. These spillovers are statistically significant at the 5%, and their significance is robust to different measurements and functional forms for inventor productivity and time aggregation.

In the second part of the paper, to quantify the importance of these spillovers to implement R&D policies, I build a quantitative model of innovation. In each location, a representative firm produces a final good that is consumed locally and is produced by aggregating intermediate inputs from all locations. Each intermediate input is produced by a single firm in each location. Firms hire workers that produce the input, and inventors that engage in R&D. In the model, R&D determines the quality of an input, where firm's inventors generate ideas heterogeneous in productivity, which are then implemented into the firm's blueprint to produce the input at a given quality.

Then, each firm optimally decides how many workers and inventors to hire subject to the demand of its input and its quality. When the firm decides how many inventors to hire, I show that the quality of an input is comprised by the number of ideas a firm's inventors generated, and by how productive these ideas are in expectation. For the first part of input quality, I assume decreasing returns to R&D, so only a subset of firm's inventors generate ideas. This is a valid and necessary assumption since I estimate it in the data, and it is the congestion force that rules out an equilibrium where all inventors move to a single location. For the second part of input quality, following the evidence on local knowledge spillovers in R&D and distributional assumptions on the process on how inventors generate ideas, I show that the expected productivity of firms inventors' ideas increases with the local knowledge spillovers in R&D in a location.

Finally, I also allow for labor mobility, so workers and inventors choose where to work according to real wages, amenities, and migration costs. And finally, the model allows for straightforward aggregation where aggregate productivity is endogenously determined in general equilibrium. The main prediction of the model is that a location's productivity is endogenously determined by three forces. First, locations with better production fundamentals or that hold more inventors are more productive due to local knowledge spillovers in R&D. Second, locations that exhibit higher labor costs are less productive since firms are less able to hire inventors to innovate. Third, locations with higher market access are more productive since higher demand from other locations increases firms' profitability, and therefore their incentive to invest in R&D. All these forces shape location's productivity in general equilibrium. Additionally, the model predicts that a location's productivity acts as an agglomeration force for overall economic activity. Since a location's productivity is determined by the its number of inventors, then locations with more inventors exhibit larger shares in locations' expenditure of intermediate inputs.

In the third part of the paper, I calibrate the model and use it to conduct policy counterfactuals and quantify the importance of local knowledge spillovers in R&D for aggregate productivity. I now describe how I discipline the model. First, the model generates an expression that establishes a relationship between inventor productivity and cluster size. This expression is the model counterpart of the specification I used to causally estimate local knowledge spillovers in R&D in the data. Then, I can directly import the estimated spillovers into the model. Second, I estimate firm-level decreasing returns to R&D by regressing the number of firm's inventors that filed a patent against the number of hired inventors by the firm. I find an elasticity of 0.65, which confirms the existence of firm-level decreasing returns to R&D. Third, I calibrate migration costs by targeting overall migration rates and estimating migration cost elasticities for both workers and inventors. Finally, I follow Redding (2016) and use aggregate data on wages and the number of workers and inventors across locations to recover unobserved fundamental location productivities and amenities.

After calibrating the model, I conduct counterfactuals to quantify the effect of policies that promote R&D activities on aggregate productivity, and the importance of local knowledge spillovers in R&D for the effectiveness of these policies. First, I simulate a supply-side policy of reducing inventor migration costs by 25%. I find that this reduction leads to a 5.87% increase in aggregate productivity. Since the total number of inventors is finite, the policy exhibits substantial heterogeneous effects across locations. I find that the increase in aggregate productivity arises from inventors moving from larger towards smaller clusters in pursue of higher real wages, so the policy reduces the spatial concentration of inventors. Second, I simulate a demand-side policy of a 25% subsidy for firms' R&D expenditure from the 2020 German R&D Tax Allowance Act. I find that this subsidy leads to a 4.27% increase in aggregate productivity. In contrast the reduction of inventor migration costs, all locations increase their productivity and the spatial concentration of inventors increases, so larger clusters exhibit higher productivity gains. Finally, I show that local knowledge spillovers in R&D are important for the effectiveness of these policies to foster aggregate productivity.

Literature. This paper contributes to three literature strands. First, this paper contributes to the empirical literature on local knowledge spillovers (Griliches, 1991; Jaffe et al., 1993; Audretsch and Feldman, 1996; Jaffe et al., 2000; Thompson, 2006; Carlino et al., 2007; Combes et al., 2010; Greenstone et al., 2010; Bloom et al., 2013; Kerr and Kominers, 2015; Kantor and Whalley, 2019; Moretti, 2021; Gruber et al., 2022). This literature largely focuses on the agglomeration of economic activity, and the positive externalities arising from it. More recently, Moretti (2021) focused in R&D and estimated local knowledge spillovers for inventors. I contribute to this literature by exploiting a historical natural experiment to causally estimate local knowledge spillovers in R&D.

Second, this paper contributes to the literature on the importance of knowledge spillovers for innovation. This is a vast literature with contributions from urban economics (Eaton and Eckstein, 1997; Glaeser, 1999; Black and Henderson, 1999; Kelly and Hageman, 1999; Duranton and Puga, 2001; Duranton, 2007; Roca and Puga, 2017; Duranton and Puga, 2019; Davis and Dingel, 2019), trade (Ramondo et al., 2016; Hallak and Sivadasan, 2013; Atkeson and Burstein, 2010; Melitz, 2003; Eaton and Kortum, 2002; Krugman, 1980; Akcigit et al., 2021), and spatial economics Desmet and Rossi-Hansberg (2014); Desmet et al. (2018); Nagy et al. (2016); Mestieri et al. (2021); Williams (2023); Crews (2023). I contribute to this literature by building a quantitative framework that explicitly accounts for local knowledge spillovers in R&D I estimate in the data.

Third, this paper contributes to the literature on policies that promote economic growth. On one side, labor mobility can have aggregate implications both in the data (Borjas and Doran, 2012; Burchardi and Hassan, 2013; Moser et al., 2014; Peri et al., 2015; Bosetti et al., 2015; Bahar et al., 2020; Burchardi et al., 2020) and in quantitative settings Monras (2018); Bryan and Morten (2019); Peters (2022); Arkolakis et al. (2020); Jaworski et al. (2020); Pellegrina and Sotelo (2021); Prato (2021). I contribute by showing the aggregate importance of reducing of migration costs for inventors. This is a significant departure from previous quantitative frameworks which do not focus on inventors explicitly.[2] On the other side, R&D policies can promote aggregate innovation Goolsbee (1998); Romer (2000); Wilson (2009); Acemoglu et al. (2018); Koehler (2018); Akcigit et al. (2021). I contribute by providing a framework that allows policy makers to evaluate the aggregate implications of R&D implications in general equilibrium.

The remainder of this paper is structured as follows. Section 1 explains how I estimate local knowledge spillovers in R&D. Section 2 describes the model. Section 3 maps the model to the data. Section 4 presents the results of the counterfactuals. Section 5 concludes.

1 Local Knowledge Spillovers in R&D

In this section I describe the estimation of local knowledge spillovers in R&D. The first part of this section describes the data, the second part explains the estimation strategy, and the third part discusses assumptions and results throughout this section. Appendices A-B contain additional tables and figures. Further details about the data are in the Supplementary Appendix.

[2]A notable exception is Koike-Mori et al. (2023)

1.1 Data sources

Linked Inventor Biography (INV-BIO). The main dataset in this paper is the INV-BIO by the Research Data Centre of the German Federal Employment Agency at the Institute for Employment Research (FDZ-IAB). The INV-BIO is an administrative dataset comprised by approximately $150,000$ German inventors with high–frequency and detailed information on their employment spells and patenting activities between 1980 and 2014. The INV-BIO is comprised by three modules: (i) an inventor-level module that includes data on inventors' job spells; (ii) an establishment-level module with yearly characteristics of inventors' establishments; and (iii) a patent-level module with information on German inventors' patents.

Sample of Integrated Employer-Employee Data (SIEED). The FDZ-IAB's SIEED is a 1.5% sample of all establishments in Germany between 1975 and 2018. The dataset tracks establishments' characteristics over time, and establishments' employees' spells over the entire period. I use this complementary dataset to compare the spatial concentration of workers to inventors, and to construct aggregate variables I later use to estimate the model.

1.2 Construction of variables.

Dimensions. From the INV-BIO modules I construct an unbalanced panel dataset of inventors. An observation in the data is an inventor i working for establishment ω in location d in technological area a during period t. I focus my analysis on West Germany, which is comprised by 104 labor markets. A labor market is defined based on commuting patterns between districts (Kosfeld and Werner, 2012), and are the equivalent to US commuting zones. Finally, to estimate long run estimates of local knowledge spillovers in R&D, I stack the data in three 10-year periods: (i) 1982-1991, (ii) 1992-2001, and (iii) 2002-2011.[3]

[3] I also consider six 5-year periods to estimate shorter run local knowledge spillovers in R&D: (i) 1982-1986, (ii) 1987-1991, (iii) 1992-1996, (iv) 1997-2001, (v) 2002-2006, and (vi) 2007-2011.

West German technological clusters. I define a technological cluster as a technological area-location pair. For example, "Mechanical engineering" in "Munich" is a cluster in West Germany. There are 5 technological areas in the data: (i) Electrical engineering, (ii) Instruments, (iii) Chemistry, (iv) Mechanical Engineering, and (v) Others. Then, locations and technological areas comprise $104 \times 5 = 520 \ (d, a)$ technological clusters.

Inventor's cluster. To define an inventor's cluster at a given period, it is necessary to determine the inventor's location and the technological area the inventor works in. First, the location of an inventor is determined by the location of the inventor's establishment since knowledge spillovers arguably happen mostly at the workplace. Additionally, since I consider establishments and not multi-location firms, the location of the inventor is unique. Second, an inventor belongs to the technological area for which he filed the highest share of patents during a given period. For example, if between 1982 and 1991, an inventor filed 80% of his patents in Chemistry, then he belongs to that technological area.

A data limitation is that inventors do not necessarily file patents every period. This generates sample selection, since only inventors that filed a patent during a given period are registered in the data. The main problem arising from this limitation is that it is not straightforward to assign a cluster to an inventor that did not file a patent during a given period. To address this problem, if an inventor did not file a patent during a given period, I assume that an inventor's cluster did not change since since the last time an inventor filed a patent. For example, if in 1995 the latest patent an inventor filed was a Chemistry patent in Dusseldorf in 1993, then I assume that in 1994-1995 the inventor kept working in the Chemistry/Dusseldorf cluster. This is a safe assumption since establishments rarely change locations and inventors tend to specialize in technological areas.

Inventor productivity and cluster size. To test for local knowledge spillovers in R&D, I construct two main variables. First, I measure inventor productivity $Z^{i\omega}_{da,t}$ as the total number of 5-year forward citations of inventor i's filled patents during period t by the German Patent and Trade Mark Office (DPMA, due to its name in German). If an inventor did not file a patent during period t, then $Z^{i\omega}_{da,t} = 0$. Second, I measure cluster size $R_{da,t}$ as the number of inventors working in cluster (d, a) at the end of period t.

Additional variables. I construct four additional variables I use for both the estimation of local knowledge spillovers in R&D in Section 1.3 and model calibration in Section 3. First, I measure the distance between every location pair $dist_{od}$ as the Euclidean distance (in miles) between the centroids of every labor market in Germany. The district maps were downloaded from the Federal Agency for Cartography and Geodesy, and the correspondence between districts and labor markets is given by Kosfeld and Werner (2012). Second, I measure the technological composition of every location, $TechComp_{da}$, by calculating location d's share of filed patents in technological area a such that $\sum_a TechComp_{da} = 1, \forall d$. Third, I measure migration shares during a given period between every location pair $\left\{\eta^L_{od,t}, \eta^R_{od,t}\right\}$ for workers and inventors, respectively. Fourth, I measure average wages in a given period for every location $\left\{w^L_{o,t}, w^R_{o,t}\right\}$ for workers and inventors, respectively.

1.3 Estimation

1.3.1 OLS estimates

To measure local knowledge spillovers in R&D, I consider the following specification between inventor productivity $Z^{i\omega}_{da,t}$ and cluster size $R_{da,t}$:

$$\log\left(Z^{i\omega}_{da,t}\right) = \iota_{d,t} + \iota_{a,t} + \iota_{da} + \iota_{\omega} + \iota_i + \beta \log\left(R_{da,t}\right) + \epsilon^{i\omega}_{da,t}. \tag{1}$$

If there are local knowledge spillovers in R&D, then $\beta > 0$. I saturate the model with a large set of fixed effects. $\iota_{d,t}$ are location/period fixed effects that account for amenities and location shocks that drive the overall activity of a location. $\iota_{a,t}$ are technological area/period fixed effects that account for overall technological shocks. ι_{da} are cluster fixed effects that account for time-invariant cluster productivity, and for the fact that some clusters file more patents than others in average. ι_ω are establishment fixed effects that account for inventor sorting due to time-invariante establishment productivity. ι_i are inventor fixed effects that control for inventor sorting due to time-invariant inventor productivity. In all specifications, standard errors are clustered at the (d, a) level. The identification assumption is that inventor unobservables $\epsilon^{i\omega}_{da,t}$ are uncorrelated with cluster size $R_{da,t}$.

The main measurement challenge is to account for zeros in the dependent variable $Z^{i\omega}_{da,t}$. I consider $\log\left(1 + Z^{i\omega}_{da,t}\right)$ as the dependent variable for the main specifications. Table 1 report the OLS estimates of Equation (1). Columns (1) − (6) show the value of the estimated spillovers as I progressively include the aforementioned fixed effects. The value of these estimates remain around 0.12. Column (6) reports the main OLS estimate that includes inventors fixed effects, which is key to compare a given inventor across periods and clusters. This estimate indicates that an inventor whose cluster size increased by 10% or moved to a cluster with 10% more inventors reports productivity gains of 1.75% in average.

Table 1: OLS models

	(1)	(2)	(3)	(4)	(5)	(6)
$\log(R_{da,t})$	0.0705	0.111	0.0985	0.109	0.0896	0.175
	(0.0256)	(0.0170)	(0.0166)	(0.0385)	(0.0358)	(0.0660)
$\iota_{d,t}$		✓	✓	✓	✓	✓
$\iota_{a,t}$			✓	✓	✓	✓
ι_{da}				✓	✓	✓
ι_ω					✓	✓
ι_i						✓
N	177,301	177,300	177,300	177,294	162,803	84,639
R^2	0.008	0.053	0.064	0.079	0.246	0.700

Notes: In this Table I report OLS estimates from Equation (1). The dependent variable is measured as $\log\left(1 + Z^{i\omega}_{da,t}\right)$, and $Z^{i\omega}_{da,t}$ is the number of 5-year forward citations from the DPMA. The table is comprised by 6 columns. Each column corresponds to a different combination of fixed effects, as pointed out by rows $4-8$. Row 2 reports the estimate of β. Row 3 reports standard errors clustered at the (d, a) level. Rows $9-10$ report the number of observations and the goodness of fit, respectively.

Robustness. Table A.2 contains the estimated spillovers under different specifications of inventor productivity. Since column (6) is the main specification in Table 1 I focus the robustness discussion around this specification. Panel A shows results when patent citations arose from the European Patent Office (EPO) and the EU (both the DPMA and EPO), respectively. Under these specifications, column (6) shows that the elasticities of inventor productivity to cluster size are 0.173 and 0.245, respectively. These spillovers are comparable to the ones reported in Table 1. Panel B shows results when I account for zeros by using the Inverse Hyperbolic Sine (IHS) for inventor citations instead of $\log(1+Z)$. Column (6) shows that the elasticity of inventor productivity to cluster size is 0.217, which is similar to the baseline estimate of 0.175. Additionally, when patent citations arose from the EPO and the EU, elasticities are around $0.21 - 0.24$.

Finally, results also hold under shorter time horizons. In Table A.3, I show the estimated spillovers when the frequency of the data is 5-year periods, where the first row measures inventor productivity as $\log(1+Z)$, and the third row measures it as $IHS(x)$. In both cases, column (6) shows that the spillovers are around 0.1, so the magnitude of local knowledge spillovers in R&D scale with the frequency of the data. The intuition for these results is that longer time horizons allow for larger spillovers to manifest in the data.

1.3.2 IV approach

Endogeneity concerns. To causally estimate β from Equation (1), the key identification assumption is that the unobservables $\epsilon_{da,t}^{i\omega}$ are uncorrelated to cluster size $R_{da,t}$. Nevertheless, there are at least two endogeneity concerns that could potentially violate this assumption. First, unobserved time-varying idiosyncratic shocks can bias the estimate of β. For example, inventors can decide to start working in a given technological area due to unobservable reasons. If inventors at the beginning of their careers in a given technology, who initially report low productivity, move to large clusters due to better career prospects, this would introduce a downward bias on β. On the other side, if inventors at the peak of their careers in a given technology, who report high productivity, move to large clusters due to even better career prospects, this would introduce an upward bias on β.

Second, unobserved time-varying cluster-level shocks can introduce an upward bias when estimating β. For example, a sudden increase in growth expectations for Chemistry in Dusseldorf could increase both cluster size due to an inflow of inventors towards that cluster, and inventor productivity in that cluster, introducing an upward bias on β. Finally, measurement error could also bias the estimate of β downwards. To address these endogeneity concerns, I then propose an instrumental variable approach to causally estimate β. In summary, I leverage quasi-exogenous variation in cluster size arising from the arrival of East German inventors towards West German clusters during the Reunification of Germany in 1990.

Brief historical background: The Reunification of Germany. During the final phase of World War II, the Potsdam Agreement was signed between the US, the UK, and the USSR on August 1st 1945. Part of this agreement was the division of Germany in two main blocs: (i) the Federal Republic of Germany (FRG, also known as "West Germany"), and (ii) the German Democratic Republic (GDR, also known as "East Germany"). FRG was based on liberal economic-social institutions from the West, while GDR was based on socialist institutions from the ex-Soviet Union.

In 1952, the borders between East and West Germany were well-established. Nevertheless, migration was still allowed between the two blocs. This lasted until 1961, when migration between these two blocs ceased. Then, in October 3rd 1990, the GDR was dissolved and the process to reunify Germany began. During this period, the "Exodus to the West" started, where a large number of East Germans migrated to the West. Figure A.2 plots the magnitude of this shock, which was considered to be unexpected and be permanent at the time. Since inventors from East Germany also moved to the West (Hoisl et al., 2016), I use the variation arising from the arrival of East German inventors across West German clusters.

IV estimates. To motivate the design of my instrument, consider an ideal experiment to causally estimate local knowledge spillovers in R&D. In this thought experiment, I would randomize inventors' clusters in West Germany, such that productivity gains arising from changes in cluster size can be estimated. Since it is not possible to obtain such exogenous variation, I extract quasi-exogenous variation in cluster size from the Reunification of Germany. To do this, I construct a shift-share instrument based on the arrival of East German inventors across West German clusters. If the variation in cluster size arising from the overall arrival of East German inventors is as-good-as-random, then this is sufficient to causally estimate local knowledge spillovers in R&D. First, I use variation in the arrival of inventors towards West German clusters, so the second stage regression in first differences of Equation (1) is

$$\Delta \log \left(Z_{da,t}^{i\omega}\right) = \iota_{d,t} + \iota_{a,t} + \beta \Delta \log \left(R_{da,t}\right) + \Delta \epsilon_{da,t}^{i\omega}. \tag{2}$$

Notice that the fixed effects in Equation (2) that prevail after introducing first-differences are location/period $\iota_{d,t}$ and technological area/period $\iota_{a,t}$ fixed effects. $\iota_{d,t}$ are crucial to control for the overall arrival of East Germans to West German locations during the Reunification. Also, $\iota_{a,t}$ accounts for overall technological change that could have happened during Reunification. Now, the first stage regression is a shift-share instrument:

$$IV_{da,t} = \sum_{o \in \mathcal{E}} g_{o,t} \times s_{o,da}, \tag{3}$$

where $o \in \mathcal{E}$ is location o in East Germany ($\mathcal{E}$), and d is a location in West Germany. The instrument is constructed as the interaction of two terms: (i) a common set of shocks to West German clusters $g_{o,t}$ (i.e. the "shifts"); and (ii) a set of exposure weights to these shocks $s_{o,da}$ (i.e. the "shares"). The shifts $g_{o,t} \equiv \log\left(\Delta R_{o,t}^{-d,-a}\right)$ are the log of the number of inventors in o that moved to any West German cluster except the instrumented cluster (d, a) during period t. Following Borusyak et al. (2022), the identification assumption to estimate β is that the overall arrival of East German inventors in West Germany $g_{o,t}$ excluding the instrumented cluster is as-good-as-random. That is, the shifts are uncorrelated with inventor unobservables within the instrumented cluster. This is a safe assumption since the instrumented cluster is being left out to construct each shift, so the shifts are constructed based solely on push factors arising from each East German location, and are clean from pull factors coming from the instrumented cluster.

The shifts are then weighted by exposure shares, which help predicting how many inventors from each East German cluster will move to each West German cluster. The shares $s_{o,da} \equiv dist_{o,d}^{-1} \times TechComp_{o,a}$ are comprised by two terms: (i) $dist_{o,d}^{-1}$ is the inverse distance between o and d; and (ii) $TechComp_{o,a}$ is the technological composition of location o. The construction of these variables is detailed in Section 1.2. The intuition of the shares is the following. First, migration flows decay with distance, so locations closer to each other should exhibit higher migration shares. This is consistent with Hoisl et al. (2016) who find that distance was indeed a key predictor for the migration from the East to the West. Second, the specialization of East German locations towards different technologies predicting which technological area an East German inventor will work on upon moving to the West. The shares are then normalized such that $\sum_{o \in \mathcal{E}} s_{o,da} = 1, \forall d, a$.

Table 2 contains the IV estimates of local knowledge spillovers in R&D. All the estimates exhibit an F-statistic above 10, which reflects the relevance of the proposed instrument. Column (1) reports the estimate of the spillovers when I do not consider any fixed effects. This reports a value of 0.178 which is similar to the OLS estimate from column (6) in Table (1). It is crucial to include location-period fixed effects to account for the overall arrival of East Germans to West Germany. In column (2) I show that estimated spillovers after including these fixed effects are 0.309. Finally, it is also key to include technological area-period fixed effects to control for technological changes after the Reunification. Column (3) contains the main empirical result of this paper: an inventor whose cluster size increases by 10% or moved to a cluster with 10% more inventors becomes 4.09% more productive in average. This estimate is between 2 and 3 times the OLS estimate of 1.75% from column (6) in Table (1), which reflects a downward bias when estimating β due to unobservables and measurement error.

Table 2: IV models

	(1)	(2)	(3)
$\Delta \log (R_{da,t})$	0.178	0.309	0.409
	(0.0431)	(0.101)	(0.152)
$\iota_{d,t}$		✓	✓
$\iota_{a,t}$			✓
$KP-F$	132.1	34.14	28.23
N	$50,778$	$50,776$	$50,776$

Notes: In this Table I report IV estimates from Equation (2), where the instrument is constructed as in Equation (3). The dependent variable is measured as $\Delta \log \left(1 + Z^{i\omega}_{da,t}\right)$, and $Z^{i\omega}_{da,t}$ is the number of 5-year forward citations from the DPMA. The table is comprised by 4 columns. Each column corresponds to a different combination of fixed effects, as pointed out by rows $5-6$. The fourth column reports the OLS estimate from Equation (2). Row 3 reports the estimate of β. Row 4 reports standard errors clustered at the (d, a) level. Rows $7-8$ report the first stage Kleibergen-Paap F-statistic (KP-F) and the number of observations, respectively.

Robustness. Table A.4 contains the estimated spillovers under different specifications of inventor productivity. Since column (3) is the main specification in Table 2, I focus the robustness discussion around this specification. Panel A shows results when patent citations arose from the EPO and the EU, respectively. Under these specifications, column (3) shows that the elasticities of inventor productivity to cluster size are 0.209 and 0.343, respectively. These spillovers are comparable but somewhat lower to the ones reported in Table 2. Panel B shows results when I account for zeros by using the IHS for inventor citations instead of $\log (1 + Z)$. Column (3) shows that the elasticity of inventor productivity to cluster size is 0.498. Additionally, when patent citations arose from the EPO and the EU, elasticities are around $0.23 - 0.39$.

Finally, results also hold under shorter time horizons. In Table A.5 I show the estimated spillovers when the frequency of the data is 5-year periods, where the first row measures inventor productivity as $\Delta \log (1 + Z)$, and the third row measures it as $\Delta IHS (x)$. In both cases, column (3) shows that the spillovers are around 0.09, so the magnitude of local knowledge spillovers in R&D scale with the frequency of the data. The intuition for these results is that longer time horizons allow for larger spillovers to manifest in the data.

1.4 Discussions

Do citations measure productivity? Throughout this paper, I have measured inventor productivity as the number of forward citations of all inventor's filed patents during a given period. Then, it is reasonable to pose whether number of citations indeed measure productivity. There is a vast literature that documents a positive relationship between number of citations and proxies for productivity, such as patent value (Kogan et al., 2017; Hall et al., 2001; Harhoff et al., 1999; Trajtenberg, 1990).

More recently, Abrams et al. (2013) find preliminary evidence of a inverse U-shaped relationship between number of citations and patent value in the data. They rationalize this finding by distinguishing between productive and strategic patents. For the former, more citations reflect a higher patent productivity since a citation reflects further creation of patents. For the latter, patenting an idea maintain incumbent's monopoly power such that entry is inhibited, so the number of citations decreases. To check whether German citations are mostly productive or strategic, I review literature on firm surveys about their incentives to patent (Blind et al., 2006; Cohen et al., 2002; Pitkethly, 2001; Duguet and Kabla, 2000; Schalk et al., 1999; Arundel et al., 1995), which is mostly focused on Europe, particularly Germany. In general, the major motive for German firms to file patents is the classical incentive to protect their ideas, which goes in line with productive patenting.

Is it exposure instead of knowledge spillovers? A possible identification threat to estimate β is that the number of citations reflect higher exposure of an inventor's ideas, which is orthogonal to knowledge spillovers. For example, if an inventor moves to a larger cluster, then his ideas could obtain more exposure to a larger share of inventors, so his patents get cited more often. This would introduce an upward bias when estimating β. I present two main arguments against this concern.

First, the patenting market is drastically different from other industries that rely on citations, such as academia. In academia, citations measure aspects other than productivity such as reputation, exposure, among others. In the patenting market, citations are required whenever an invention uses information from another patent. Whenever a citation this situation does not take place, a patent infringement has taken place, so then the owner of the non-cited patent can pursue legal means to resolve the issue. This is particularly relevant for the industrial economy of Germany that reports one of the largest number of patent litigation cases (Cremers et al., 2017), and exhibits one of the highest cross-country levels of patent enforcement (Papageorgiadis and Sofka, 2020).

Second, assuming that these effects are biasing the estimate of β, Tables A.2 and A.4 include the OLS and IV estimates where productivity is measured by the number of citations from the EPO, which is the European patenting institution and completely independent from the German patenting office. These estimates still provide evidence on the existence of local knowledge spillovers in R&D.

Comparison to previous estimates. I now compare my estimates with previous literature. Carlino et al. (2007) shows that the rate of patenting per capita is around 1.95% higher in a US metropolitan area with 10% higher population density. My baseline estimate of 4.09% is higher due to three differences. First, I test for knowledge spillovers in R&D by measuring productivity through number of citations instead of patenting rates. Second, I estimate long-run local knowledge spillovers since I consider 10-year periods. In contrast, they leverage cross-sectional variation across US metropolitan areas. Third, my identification relies on a historical natural experiment instead on the inclusion of covariates.

Moretti (2021) is the closest to this paper. His OLS estimate is around 0.067, while my estimate from Table 1 is 0.175. When running the model in first differences, his IV estimate is around 0.049, while my estimates from Table 2 is 0.409. Even thought both of these papers estimate local knowledge spillovers in R&D at the inventor level, the differences in magnitudes arise due to two differences. First, I estimate long-run spillovers (10-year periods), while Moretti estimates short-run spillovers (1-year periods). Second, my larger estimates could result from stronger local knowledge spillovers in R&D in Germany in comparison to the US.

2 Model

In this section I build a quantitative spatial model of innovation. The Supplementary Appendix contains detailed derivations, microfoundations, and solution algorithms for the model.

2.1 Setup

Geography. There is a discrete set of locations $\mathcal{S} \equiv \{1, 2, \ldots, S\}$, where $o \in \mathcal{S}$ is the origin location, and $d \in \mathcal{S}$ is the destination location.

Firms. There are two types of firms in each location: (i) a final good firm, and (ii) an intermediate input firm. The final good is produced by a representative firm, it is non-tradable, and it is produced by aggregating intermediate inputs from all locations with constant elasticity of substitution (CES). The intermediate input firm produces a representative intermediate input, which is tradable across locations subject to trade costs. This firm is comprised by a production facility and a colocated R&D subsidiary. The R&D subsidiary freely transmits a blueprint to the facility, which the facility uses to produce the firm's intermediate input at a given quality. The production facility hires workers, and the R&D subsidiary hires inventors.

Agents. Each agent is either an inventor or a worker. Each agent supplies a unit of labor inelastically, earns income from wages, housing rent, and redistributed profits, and consumes local final goods. Agents are mobile, so they optimally decide where to locate by maximizing their utility subject to migration costs.

2.2 Technology

Final goods. In each location d, a representative firm produces a final good by aggregating intermediate inputs from all locations. The production function of the final good is

$$Q_d = \left(\sum_o Z_o^{\frac{1}{\sigma}} Q_{od}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{4}$$

where Q_d is the production of the final good, Q_{od} is the quantity of intermediate inputs from o sold to the final good firm in d, Z_o is the quality of the intermediate input, and $\sigma > 1$ is the constant elasticity of substitution (CES) across intermediate inputs. The final good firm maximizes profits subject to Equation (4), which yields the demand for intermediate inputs

$$Q_{od} = Z_o P_{od}^{-\sigma} P_d^{\sigma-1} X_d, \tag{5}$$

where $P_d^{1-\sigma} = \sum_o Z_o P_{od}^{1-\sigma}$ is CES price index, and $X_d \equiv P_d Q_d$ is total expenditure on the final good in d. From Equation (5), quality Z_o acts as a demand shifter for the intermediate input o.

Production of intermediate inputs. In each location o, there is a representitative producer of a tradable intermediate input. Each firm owns a production facility and a colocated R&D center. The R&D center freely shares a blueprint with the facility that describes how to produce the firm's intermediate input at quality Z_o. Then, given input quality Z_o, the firm optimally produces its intermediate input. The profits of the firm in o selling to d is

$$\pi_{od} = P_{od} Q_{od} - w_o^L L_{od}, \tag{6}$$

where L_{od} is the demand for workers. A unit of labor is required to produce an intermediate input, so

$$L_{od} = \frac{\tau_{od} Q_{od}}{\mathcal{A}_o}, \tag{7}$$

where $\mathcal{A}_o$ are location fundamentals for the production of intermediate inputs, w_o^L are worker wages, L_{od} is the demand for workers, and $\tau_{od} > 1$ are iceberg trade costs. Then, the firm maximizes total profits $\pi_o = \sum_d \pi_{od}$ subject to Equations (5), (6), and (7). Optimally, the firm charges a constant markup on its unit cost

$$P_{od} = \overline{m} \frac{\tau_{od} w_o^L}{\mathcal{A}_o}, \tag{8}$$

where $\overline{m} \equiv \frac{\sigma}{\sigma - 1}$ is the CES constant markup over marginal costs. By plugging back Equation (8) into firm total profits, we can rewrite them as

$$\pi_o = \frac{1}{\sigma} Z_o \sum_d \left(\frac{P_{od}}{P_d} \right)^{1-\sigma} X_d. \tag{9}$$

From Equation (9), total profits of the firm in o increases proportionally with the quality of its intermediate input Z_o. This reflects the role of quality Z_o acting as a demand shift for the intermediate input in Equation (5).

Quality of intermediate inputs. The R&D subsidiary of the intermediate input firm in hires R_o inventors, and each inventor produces an idea which is then implemented into a blueprint that describes how to produce the firm's intermediate input. Then, the quality of the intermediate input is

$$Z_o = \mathbb{Z}_o R_o, \tag{10}$$

where $\mathbb{Z}_o$ is the expected productivity of inventors' ideas. In the Supplementary Appendix, I provide two microfoundations that generate isomorphic expressions for the quality of intermediate inputs, up to a constant. To provide intuition on how the quality of the intermediate input is determined by the expected productivity of inventors' ideas, I briefly sketch the first microfoundation based on necessary tasks.

Consider that the firm's R&D subsidiary owns a blueprint that contains a continuum of tasks to produce a unit the firm's intermediate input. The R&D center hires R_o inventors where each one produces an idea to be implemented in the blueprint, and these ideas are heterogeneous in productivity. Given the assumption that each of these ideas improve the quality of every task within the blueprint, and all tasks are necessary to produce a unit of the intermediate input, then the expected productivity of the implemented ideas into the blueprint captures the overall quality of the firm's intermediate input.

Productivity of an inventor's ideas. An inventor i working at a R&D subsidiary generates an idea to be implemented into the firm's blueprint on how to produce a unit of the firm's input. Ideas are heterogeneous in productivity Z_o^i drawn from a probability distribution:

$$Z_o^i \sim Frechet\left(\alpha, \lambda_o^{\frac{1}{\alpha}}\right), \tag{11}$$

where α and $\lambda_o^{\frac{1}{\alpha}}$ are the shape and scale parameters of the Frechet distribution, respectively. The Supplementary Appendix describes inventors' innovation process based on Kortum (1997) that generates a Frechet distribution for the productivity of inventors' ideas. Under this framework, λ_o is referred as the *spillover function* since it embeds exogenous economic forces that increase inventors' productivity. Considering the probability distribution in Equation (11), then the expected productivity of inventors' ideas is

$$\mathbb{Z}_o = \psi \lambda_o^{\frac{1}{\alpha}}, \tag{12}$$

where $\psi > 0$ is a constant that arises from the microfoundation for the quality of intermediate inputs. Finally, guided by the empirical evidence on local knowledge spillovers in R&D in Section 1.3, I consider the following functional form for the spillover function:

$$\lambda_o^{\frac{1}{\alpha}} = \mathcal{Z}_o R_o^{\widetilde{\gamma}}, \tag{13}$$

where $\mathcal{Z}_o$ are location fundamentals for R&D, R_o is the number of inventors in o, and $\widetilde{\gamma} \equiv \frac{\gamma}{\alpha}$ are local knowledge spillovers in R&D.[4]

Research and Development (R&D). The R&D subsidiary of the intermediate input firm engages in R&D to endogenously determine the quality of its intermediate input. The quality of the firm's input is specified by a blueprint generated by the R&D subsidiary. The blueprint is freely shared with the facility, and the firm's production facility and R&D subsidiary are colocated. Then, I can characterize the problem of firm's R&D subsidiary as a maximization of its research output Z_o after hiring inventors at wage w_o^R subject to the R&D subsidiary's production function implied by Equations (10)-(12). Then, firm demand for inventors is

$$w_o^R = \psi \mathcal{Z}_o R_o^{\widetilde{\gamma}}. \tag{14}$$

2.3 Location choice

In each location d, there are two types of agents: inventors ($n = R$), and workers ($n = L$). Upon moving to d, agents maximize their utility subject to their budget constraint. Agents have preferences for consuming local final goods, housing, and location amenities. Then, an agent's indirect utility is

$$U_d^n = \frac{\mathcal{B}_d^n V_d^n}{P_d^{\beta} r_d^{1-\beta}}, \tag{15}$$

where $\mathcal{B}_d^n$ are type-specific location amenities, $V_d^n = \frac{(1+\overline{\pi})w_d^n}{\beta}$ is income[5], r_d is housing rent, and β is the expenditure share towards final goods. Then, an agent i of type n working in o moves to d by maximizing its utility:

$$U_{od}^{i,n} = \max_{d \in \mathcal{S}} \left\{ \frac{U_d^n}{\mu_{od}^n} \times \epsilon^i \right\}, \tag{16}$$

[4] Technically, γ are local knowledge spillovers in R&D. Since γ and α are not separable, I consider $\tilde{\gamma} \equiv \frac{\gamma}{\alpha}$ to denote local knowledge spillovers in R&D throughout the paper.

[5] Housing expenditure in each location is redistributed as lump sum transfers to local workers and inventors. National firm profits are invested in a national investment fund, and redistributed uniformly among workers and inventors.

where $\mu_{od}^{n} \geq 1$ are type-specific *iceberg* migration costs, $G(\epsilon) = \exp\left(-\epsilon^{-\kappa}\right)$ are location preference shocks, and κ is the spatial labor supply elasticity. Leveraging the order-statistic properties of the Frechet distribution from Equation (16), the share of agents of type n moving from o to d is

$$\eta_{od}^{n} = \frac{\left(\frac{U_d^n}{\mu_{od}^n}\right)^{\kappa}}{\sum_{\delta}\left(\frac{U_{\delta}^n}{\mu_{o\delta}^L}\right)^{\kappa}}. \tag{17}$$

2.4 Other important variables

Quality. From Equations (10)-(12), the quality of the intermediate input in location o is

$$Z_o = \psi \mathcal{Z}_o R_o^{1+\widetilde{\gamma}}. \tag{18}$$

Price indices. From Equations (8) and (18), price indices are

$$P_{od} = \overline{m}\frac{\tau_{od} w_o^L}{\mathcal{A}_o} \quad \textit{and} \quad P_d^{1-\sigma} = \sum_o Z_o P_{od}^{1-\sigma}. \tag{19}$$

Trade shares. From Equations (19), location o's share in location d's expenditure is

$$\chi_{od} = \frac{Z_o \mathcal{A}_o^{\sigma-1}\left(\tau_{od} w_o^L\right)^{1-\sigma}}{\sum_o Z_o \mathcal{A}_o^{\sigma-1}\left(\tau_{od} w_o^L\right)^{1-\sigma}}. \tag{20}$$

Equation (20) shows that trade shares χ_{od} increase with location fundamentals for production $\mathcal{A}_o$ and quality Z_o of intermediate inputs. This highlights the role of R&D as an agglomeration force for economic activity.

2.5 Equilibrium

Trade balance. Then, total income Y_o is comprised by total wages, redistributed profits and lump sum transfers from local housing expenditure:

$$Y_o = \frac{(1+\overline{\pi})\left(w_o^L L_o + w_o^R R_o\right)}{\beta}. \tag{21}$$

Total expenditure X_o is comprised by purchased intermediates from every location d:

$$X_o = \sum_d \chi_{od} X_d. \tag{22}$$

To close the model, I impose trade balance in every location:

$$Y_o = X_o. \tag{23}$$

Housing market. From the utility maximization problem of inventors and workers, aggregate demand for housing is

$$r_o = \frac{(1-\beta)\, Y_o}{H_o}. \tag{24}$$

Aggregate supply of housing is fixed, so

$$H_o = \overline{H}_o, \tag{25}$$

where $\overline{H}_o$ is the fixed quantity of housing in location o.

Labor markets. Consider an initial distribution of workers and inventors across locations $\left\{\overline{R}_o, \overline{L}_o\right\}_{\forall o \in \mathcal{S}}$. Then, considering the migration shares from Equations (17), the aggregate supply of workers and inventors across locations are

$$R_d = \sum_o \eta_{od}^R \overline{R}_o, \tag{26}$$

$$L_d = \sum_o \eta_{od}^L \overline{L}_o. \tag{27}$$

From the demand for inventors in Equation (30) and the definition of quality in Equation (18), the demand for inventors is

$$w_o^R = \frac{Z_o}{R_o}. \tag{28}$$

Finally, since I impose trade balance, the demand for workers is not necessary to close the model due to Walras's Law.

Definition 1 (Equilibrium). Given the exogenous distribution of workers and inventors across locations $\left\{\overline{R}_o, \overline{L}_o\right\}_{\forall o \in \mathcal{S}}$, fixed supply of housing $\left\{\overline{H}_o\right\}_{\forall o \in \mathcal{S}}$, location fundamentals $\{\mathcal{Z}_o, \mathcal{A}_o\}_{\forall o \in \mathcal{S}}$, location amenities $\left\{\mathcal{B}_o^R, \mathcal{B}_o^L\right\}_{\forall o \in \mathcal{S}}$, migration costs $\left\{\mu_{od}^R, \mu_{od}^L\right\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, and parameters, *an equilibrium is a set of wages* $\left\{w_o^R, w_o^L\right\}_{\forall o \in \mathcal{S}}$*, housing rent* $\{r_o\}_{\forall o \in \mathcal{S}}$*, prices* $\{P_o\}_{\forall o \in \mathcal{S}}$*, quantities* $\{R_o, L_o, H_o, Q_o\}_{\forall o \in \mathcal{S}}$*, and quality* $\{Z_o\}_{\forall o \in \mathcal{S}}$ *such that* (i) *workers and inventors maximize utility,* (ii) *firms maximize profits,* (iii) *workers and inventors labor markets clear,* (iv) *housing markets clear, and* (v) *trade is balanced.*

2.6 Equilibrium with R&D subsidies

A national government implements location-specific subsidies for firms' expenditure in R&D $\{s_o\}_{\forall o \in \mathcal{S}}$. These subsidies are funded with a uniform labor tax τ. The government holds a balanced budget such that

$$\tau \sum_o \left(w_o^L L_o + w_o^R R_o\right) = \sum_o s_o \left(w_o^R R_o\right), \tag{29}$$

demand for inventors is

$$w_o^R = \frac{Z_o}{(1 - s_o) R_o}, \tag{30}$$

total income of a location is

$$Y_o = \frac{(1 - \tau + \overline{\pi}) \left(w_o^L L_o + w_o^R R_o \right)}{\beta}, \tag{31}$$

and housing demand is

$$r_o = \left(\frac{1 - \beta}{\beta} \right) \frac{(1 - \tau + \overline{\pi}) \left(w_o^L L_o + w_o^R R_o \right)}{H_o}. \tag{32}$$

Since the labor tax is uniform, they do not distort labor supplies, so migration shares in Equation (17) remain unchanged.

Definition 2 (Equilibrium with R&D subsidies). Given the exogenous distribution of workers and inventors across locations $\left\{\overline{R}_o, \overline{L}_o\right\}_{\forall o \in \mathcal{S}}$, fixed supply of housing $\left\{\overline{H}_o\right\}_{\forall o \in \mathcal{S}}$, location fundamentals $\{\mathcal{Z}_o, \mathcal{A}_o\}_{\forall o \in \mathcal{S}}$, location amenities $\left\{\mathcal{B}_o^R, \mathcal{B}_o^L\right\}_{\forall o \in \mathcal{S}}$, migration costs $\left\{\mu_{od}^R, \mu_{od}^L\right\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, R&D subsidies $\{s_o\}_{\forall o \in \mathcal{S}}$, and parameters*, an equilibrium with R&D subsidies is a set of wages* $\left\{w_o^R, w_o^L\right\}_{\forall o \in \mathcal{S}}$*, housing rent* $\{r_o\}_{\forall o \in \mathcal{S}}$*, prices* $\{P_o\}_{\forall o \in \mathcal{S}}$*, quantities* $\{L_o, R_o, H_o, Q_o\}_{\forall o \in \mathcal{S}}$*, and quality* $\{Z_o\}_{\forall o \in \mathcal{S}}$ *such that* (i) *workers and inventors maximize utility,* (ii) *firms maximize profits,* (iii) *workers and inventors labor markets clear,* (iv) *housing markets clear,* (v) *government's budget is balanced, and* (vi) *trade is balanced.*

3 Taking the Model to the Data

In this section I describe the calibration strategy of the model. The model is parametrized by the geography of West Germany, local knowledge spillovers in R&D $\{\widetilde{\gamma}\}$, migration costs $\left\{\mu_{od}^R, \mu_{od}^L\right\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, location fundamentals $\{\mathcal{Z}_o, \mathcal{A}_o\}_{\forall o \in \mathcal{S}}$, location amenities$\left\{\mathcal{B}_o^R, \mathcal{B}_o^L\right\}_{\forall o \in \mathcal{S}}$, trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, and remaining parameters $\{\alpha, \psi, \kappa, \sigma, \beta\}$. Table 4 at the end of this section summarizes the calibration strategy of the model. Details about the parametrization of the model are in the Supplementary Appendix.

Geography. The discrete set of locations $\mathcal{S}$ are the 104 labor markets in West Germany.

Local Knowledge Spillovers in R&D $\{\widetilde{\gamma}\}$. The reduced-form estimates for local knowledge spillovers in R&D in Section 1 are mapped to $\widetilde{\gamma}$. Consider Equation (12), which describes how cluster size increases the expected productivity of inventors. Considering Equation (13), the model yields a log-log relationship between inventor productivity and cluster size:

$$\log\left(Z_o^i\right) = \iota + \iota_o + \widetilde{\gamma}\log\left(R_o\right) + \epsilon_o^{i\omega}, \tag{33}$$

where $\iota \equiv \log(\psi)$ and $\iota_o \equiv \log(\mathcal{Z}_o)$. After considering the additional time dimension t and technological areas a, and first differences, Equation (33) is the model counterpart of Equation (2) which was used to estimate local knowledge spillovers in R&D $\beta = 0.409$. Notice that, technically, β is the elasticity of inventor 5-year forward citations to cluster size, while $\widetilde{\gamma}$ is the elasticity of patent/idea productivity or quality to cluster size. Therefore, the value of $\widetilde{\gamma}$ is such that $\widetilde{\gamma} = \delta\beta$, where δ is the elasticity of patent/idea productivity or quality to 5-year forward citations. I follow Lanjouw and Schankerman (2004) and consider $\delta = 0.22$, such that$\widetilde{\gamma} = \delta\beta = (0.22)\,(0.409) \approx 0.09$.

Migration costs $\left\{\mu_{od}^R, \mu_{od}^L\right\}$. For inventors ($n = R$) and workers ($n = L$), I parametrize migration costs as an exponential function of geographic distance between every location pair $\mu_{od}^n = \rho_0^n dist_{od}^{\rho_1^n}\exp\left(-\frac{\epsilon_{od}^n}{\kappa}\right)$, where $\{\rho_0^n\}$ are intercepts that determines the overall level of internal migration, $\{\rho_1^n\}$ are the elasticities of migration costs to distance, and ϵ_{od}^n are i.i.d. shocks. To keep the estimation consistent with the reduced-form estimates, I consider 10-year periods. I calibrate $\{\rho_0^n\}$ by targeting the 10-year average migration rates for workers and inventors of 24.99% and 26.38%, respectively. The calibrated values are

$\{\rho_0^R, \rho_0^L\} = \{1.76, 1.61\}$. To estimate $\{\rho_1^n\}$, the location choice problem of the model yields migration gravity equations for both workers and inventors:

$$\log\left(\eta_{od,t}^n\right) = \iota + \iota_{o,t} + \iota_{d,t} - \kappa\rho_1^n \log\left(dist_{od}\right) + \epsilon_{od,t}^n. \tag{34}$$

The gravity equation in (34) states that, conditional on origin/time and destination/time fixed effects $\{\iota_{o,t}, \iota_{d,t}\}$, data on geographic distance between locations, and the spatial labor supply elasticity κ, then migration elasticities to trade costs $\{\rho_1^n\}$ are identified. Since migration shares report values of zero, I estimate these elasticities through Poisson Pseudo Maximum Likelihood (PPML) estimation. From columns (2) and (4) in Table 3, I consider $\{\rho_1^R, \rho_1^L\} = \{0.602, 0.591\}$. These values are very close to the median value of migration elasticities estimated by Allen and Donaldson (2020), who also estimate them considering 10-year periods. Intuitively, Table A.6 shows that the value of these elasticities go up to $\{\rho_1^R, \rho_1^L\} = \{0.651, 0.65\}$ when considering 5-year periods, which reflect higher barriers to move in the shorter run.

Table 3: Estimation of migration costs

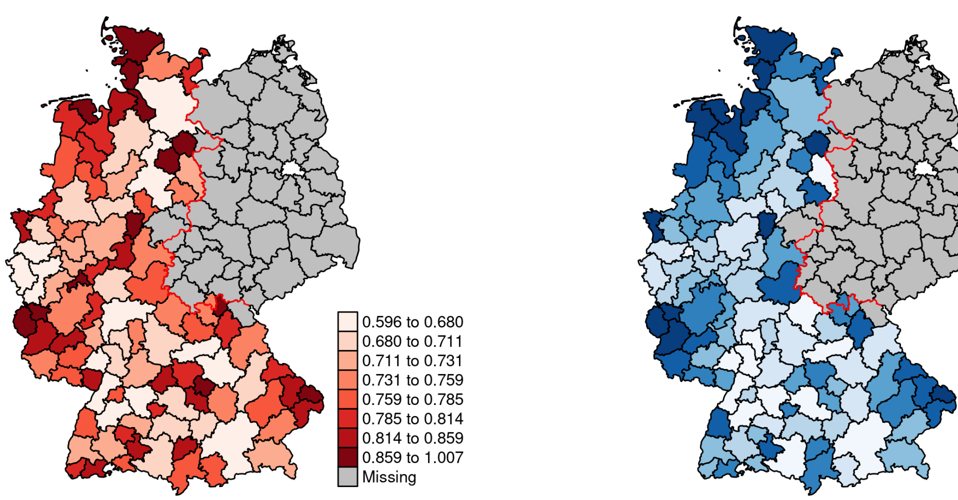
	$n = R$		$n = L$	
	OLS	$PPML$	OLS	$PPML$
$\log\left(dist_{od}\right)$	-1.001	-1.254	-1.063	-1.277
	(0.014)	(0.018)	(0.020)	(0.016)
ρ_1^n	0.472	0.591	0.501	0.602
R^2	0.812	$\cdot$	0.839	$\cdot$
N	8,336	21,632	18,381	21,632

Notes: In this table I report migration cost elasticities from Equation (34). Columns $2-3$ are the regressions for inventors, where column 2 are OLS estimates, and column 3 are PPML estimates. Columns $4-5$ are the regressions for workers, where column 4 are OLS estimates, and column 5 are PPML estimates. For OLS estimates, the dependent variable is measured as $\log\left(\eta_{od,t}^n\right)$ is the log of the share of inventors or workers from o that moved to d during a given period. Row 3 is the estimate associated to $\log\left(dist_{od}\right)$, where $dist_{od}$ is the Euclidean distance in miles from o to d. Row 4 are standard errors two-way clustered at the o, t and d, t level. Row 5 is the implied migration elasticity from the estimates from row 3 given $\kappa = 2.12$. Rows $6-7$ contain the goodness of fit and number of observations in each specification, respectively.

Location fundamentals $\{\mathcal{Z}_o, \mathcal{A}_o\}$**.** I recover location fundamentals for R&D $\{\mathcal{Z}_o\}$ and production $\{\mathcal{A}_o\}$ through model inversion. First, given parameter values $\{\psi, \widetilde{\gamma}\}$, and data on wages and population $\{w_o^R, R_o\}$, there is a unique set of values for location fundamentals for R&D $\{\mathcal{Z}_o\}$ that is consistent with the aggregate demand for inventors from Equation (28). Then, given trade costs $\{\tau_{od}\}$, location fundamentals for R&D $\{\mathcal{Z}_o\}$, parameter values $\{\psi, \sigma, \widetilde{\gamma}\}$, and data on wages and population $\{w_o^R, w_o^L, R_o, L_o\}$, there is a unique set of values for location fundamentals for production $\{\mathcal{A}_o\}$ that is consistent with trade balance from Equations (20)-(23). Since the model is static, I use data on wages and population from 2014 to denote West Germany's stationary equilibrium. In Figure 1 I show the spatial distribution of these location fundamentals. As expected, to rationalize the presence of production and innovation in less-dense locations, these locations must report higher fundamental levels of productivity.

Figure 1: Location fundamentals

(a) R&D $\{\mathcal{Z}_o\}$ (b) Production $\{\mathcal{A}_o\}$



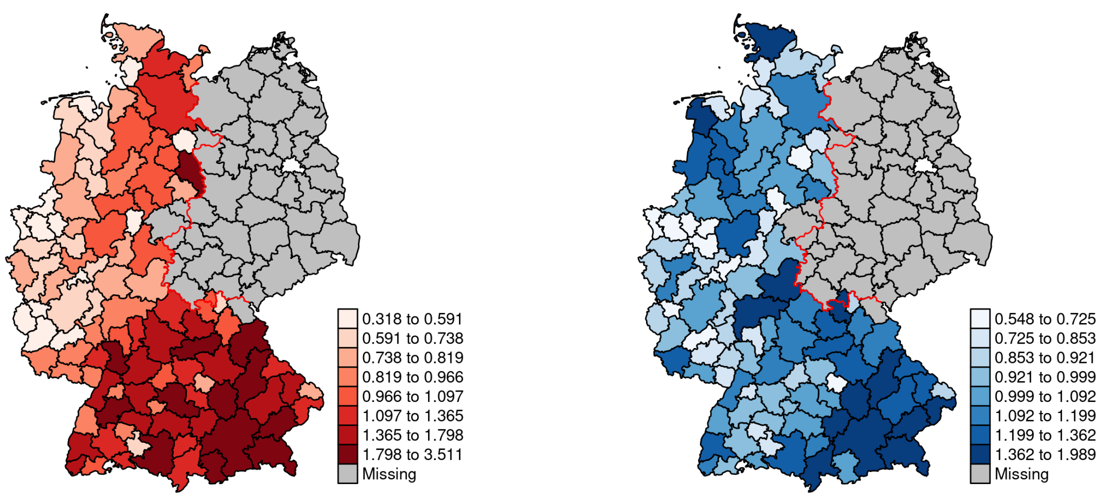
Notes: These figures show the spatial distribution of location fundamentals $\{\mathcal{Z}_o, \mathcal{A}_o\}$ in West Germany. A darker (lighter) orange color denotes a higher (lower) location fundamental. All these values are normalized by their corresponding geometric mean.

Location amenities $\{\mathcal{B}_o^R, \mathcal{B}_o^L\}$**.** I recover location amenities for both workers and inventors $\{\mathcal{B}_o^R, \mathcal{B}_o^L\}$ through model inversion. Given the exogenous distribution of workers and inventors across locations $\{\overline{R}_o, \overline{L}_o\}_{\forall o \in \mathcal{S}}$, fixed supply of housing $\{\overline{H}_o\}_{\forall o \in \mathcal{S}}$, trade costs $\{\tau_{od}\}$, migration costs $\left\{\mu_{od}^{R}, \mu_{od}^{L}\right\}$, location fundamentals $\{\mathcal{Z}_o, \mathcal{A}_o\}$, parameter values $\{\alpha, \psi, \kappa, \sigma, \beta\}$, and data on wages and population $\left\{w_o^R, w_o^L, R_o, L_o\right\}$, there is a unique set of values for location amenities $\left\{\mathcal{B}_o^R, \mathcal{B}_o^L\right\}$ that is consistent with the data. Since the model is static, I use data on wages and population from 2014 to denote West Germany' steady-state equilibrium. The initial distribution $\left\{\overline{R}_o, \overline{L}_o\right\}$ is from 1980 and they are scaled such that the total number of workers and inventors in West Germany is the same for 2014. To simplify, I consider that $\overline{H}_o = \overline{L}_o + \overline{R}_o$. To recover these amenities, I solve a fixed point algorithm on the system of excess demand functions implied by Equations (27) and (26).

In Figure 2 I show the spatial distribution of these fundamentals. As reflected by the spatial distribution of both workers and inventors, locations like Munich, Stuttgart, and Hamburg reflect the highest levels of amenities. More importantly, inventors exhibit higher levels of location amenities in the south of West Germany than workers, which reflects their higher level of spatial concentration in the data.

Figure 2: Location amenities

(a) Inventors $\left\{\mathcal{B}_o^R\right\}$ (b) Workers $\left\{\mathcal{B}_o^L\right\}$



Notes: This figure shows the spatial distribution of location amenities $\left\{\mathcal{B}_o^R, \mathcal{B}_o^L\right\}$ in West Germany. A darker (lighter) purple color denotes a higher (lower) amenity. All these values are normalized by their corresponding geometric mean.

Trade costs $\{\tau_{od}\}$. I parametrize trade costs as an exponential function of geographic distance between every location pair $\tau_{od} = \xi_0 dist_{od}^{\xi_1}$, where ξ_0 is an intercept that determines the overall level of internal trade, and ξ_1 is the elasticity of trade costs to distance. Following Ramondo et al. (2016), I calibrate ξ_0 to target a 50% share of total intra-regional trade. For the elasticity of trade costs to distance, I follow Krebs and Pflüger (2021) and set $\xi_1 = \frac{1.56}{\sigma-1}$.

Remaining parameters $\{\alpha, \psi, \kappa, \sigma, \beta\}$. The remaining parameters are the dispersion of productivity of ideas (α), the constant that arises from the microfoundation for the quality of intermediate inputs (ψ), the spatial labor supply elasticity (κ), the elasticity of substitution across intermediate inputs (σ), and the expenditure share towards final goods (β). Regardless of the microfoundation for firms' R&D, a value of α is necessary to obtain values for the constant ψ from Equation (12). Following the process for the generation of ideas, α is the Pareto shape parameter for the productivity of ideas. I run a parametric fit on the number of 5-year forward citations and set $\alpha = 1.5$. This value is similar to previous Pareto parametric fits for the number of forward citations (Silverberg and Verspagen, 2007). The value of the constant ψ depends on the microfoundation of the innovation process. Regardless of the chosen microfoundation, its value is only a function α. For the migration elasticity κ, I follow Peters (2022) and set $\kappa = 2.12$. I follow Broda and Weinstein (2006) and set $\sigma = 2.5$, which is the median elasticity for industrial sectors as in the German economy between the 1980s and 2000s. Finally, I follow Redding (2016) and set $\beta = 0.75$.

Table 4: Summary of calibration

Parameter	Description	Value	Identification/Targets
		Innovation	
$\widetilde{\gamma}$	Local Knowledge Spillovers in R&D	$\widetilde{\gamma} = (0.409)\,(0.22)$	0.409 : IV estimate, Table 2, column 3 0.22 : Lanjouw and Schankerman (2004), Table 2, column 8
α	Idea productivity dispersion	1.5	Pareto parametric fit
		Migration	
$\{\rho_0^R, \rho_0^L\}$	Migration costs, intercepts	$\rho_0^R = 1.61$ $\rho_0^L = 1.76$	26.38% migration rate of inventors 24.99% migration rate of workers
$\{\rho_1^R, \rho_1^L\}$	Migration costs, elasticities	$\rho_1^R = \frac{1.254}{\kappa}$ $\rho_1^L = \frac{1.277}{\kappa}$	Gravity estimates
κ	Migration elasticity	2.12	Peters (2022), Table 9
		Location fundamentals and amenities	
$\mathcal{Z}_o$	Location fundamentals for R&D		Recovered from aggregate demand for inventors
$\mathcal{A}_o$	Location fundamentals for production		Recovered from trade balance
$\{\mathcal{B}_o^R, \mathcal{B}_o^L\}$	Location amenities		Recovered from aggregate supply for inventors and workers
		Trade	
ξ_0	Trade costs, intercept	0.17	50% intra-trade shares (Ramondo et al., 2016)
ξ_1	Trade costs, elasticity	$\frac{1.56}{\sigma-1}$	Krebs and Pflüger (2021)
σ	Elasticity of substitution	2.5	Broda and Weinstein (2006), Table 5
β	Exp. share on final goods	0.75	Redding (2016)

Notes: This table summarizes the calibration of the model parameters. The first column shows the parameter of interest, the second column provides a short description, the third column reports the calibrated value, and the fourth column briefly describes the identification strategy.

4 R&D Policy Counterfactuals

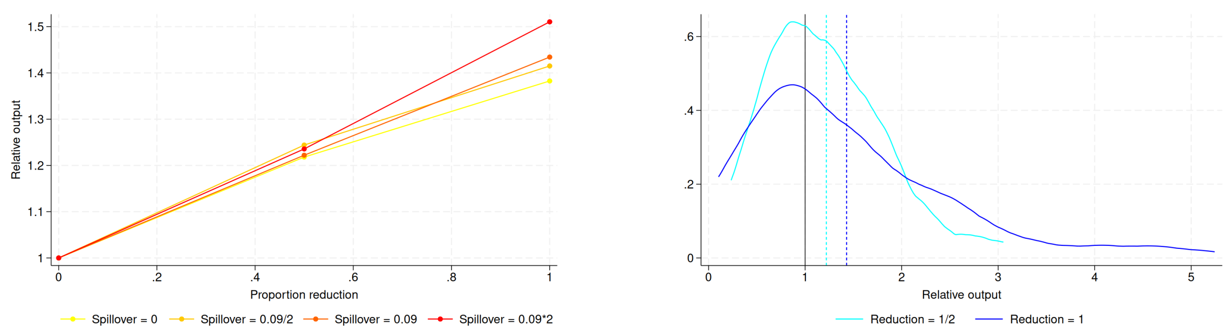
In this section, I use the calibrated model to conduct three counterfactuals to quantify the importance of local knowledge spillovers in R&D. First, I implement a reduction in migration costs for inventors. Second, I evaluate the effect of the 25% subsidy for firms' R&D expenditure within the *2020 German R&D Tax Allowance Act.*. Third, by solving the social planner problem, I back out optimal R&D subsidies..

4.1 Reducing inventor migration costs

In Figure 3, we evaluate the effect of reducing inventor migration costs μ_{od} on aggregate output Q_d. To do this, we consider different values for inventor migration costs $\widehat{\mu_{od}^R} = \mu_{od}^{R^{(1-\kappa)}}$, where $\kappa \in [0,1]$ is a proportional reduction parameter. $\kappa = 0$ is the baseline calibration for migration costs, and $\kappa = 1$ is frictionless migration for inventors. In Panel 3a, we study the aggregate effect of reducing inventor migration costs $\widehat{\mu_{od}^R}$ on aggregate output $\mathbb{E}\left\{\frac{\widehat{Q_d}}{Q_d}\right\}^{\gamma=\gamma^*}$, where $\widehat{Q_d}$ is aggregate output given $\widehat{\mu_{od}^R}$, Q_d is aggregate output given μ_{od}^R, and $\gamma = \gamma^*$ is a given value of local knowledge spillovers in R&D. We observe that, for different values of spillovers $\widetilde{\gamma}$, aggregate output significantly increases with reductions of inventor migration costs. For $\kappa = \frac{1}{2}$, output increases around 22% on average, and at $\kappa = 1$ output increases between 40% and 50% on average. Given that inventors comprise a extremely small part of the labor force, these exercises reflect the importance of promoting the geographic mobility of inventors to promote economic activity. In Panel 3b, we study the heterogeneous effects on aggregate output $\frac{\widehat{Q_d}}{Q_d}$ across West German locations for different values of $\widehat{\mu_{od}^R}$ and given the baseline value $\widetilde{\gamma} = 0.09$. We observe that the large increases in aggregate output mask a large amount of heterogeneity. For example, at a proportional reduction of $\kappa = \frac{1}{2}$, changes in output across West German locations range from -50% to 200%.

Figure 3: Reduction of inventor migration costs

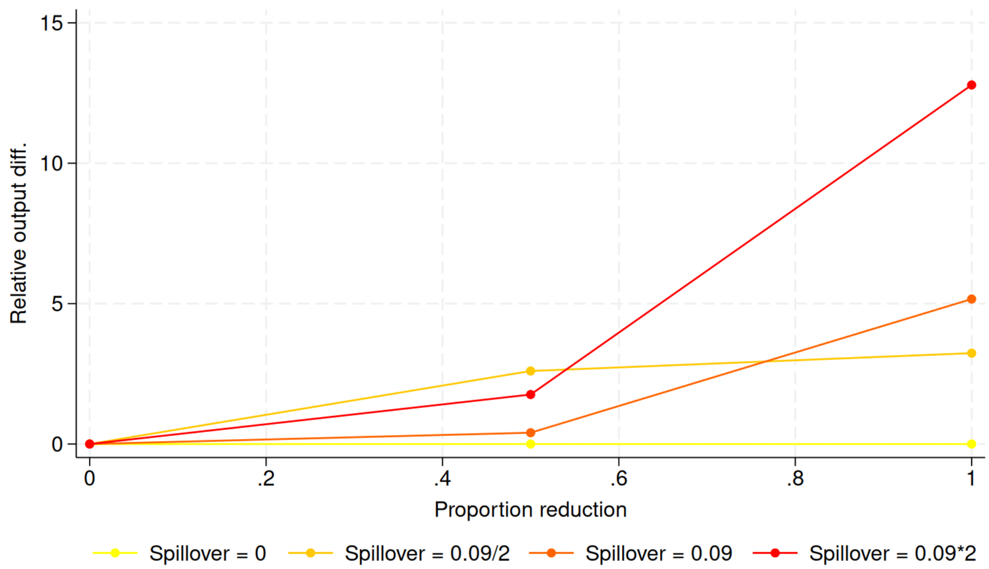
(a) Aggregate effect (b) Heterogeneous effects at $\widetilde{\gamma} = 0.09$



Notes: This figure is comprised by two panels. In Panel 3a, the horizontal axis is the proportional reduction parameter of inventor migration costs κ in $\widehat{\mu_{od}^R}$, and the vertical axis is the expected change in aggregate output $\mathbb{E}\left\{\frac{\widehat{Q_d}}{Q_d}\right\}^{\gamma=\gamma^*}$, where $\widehat{Q_d}$ is aggregate output given $\widehat{\mu_{od}^R}$, Q_d is aggregate output given μ_{od}^R, and $\widetilde{\gamma} = \gamma^*$ is a given value of local knowledge spillovers in R&D (lines yellow, orange, brown, and red). In Panel 3b we show density plots across West German locations for $\frac{\widehat{Q_d}}{Q_d}$ for proportional reductions of $\kappa = \left\{\frac{1}{2}, 1\right\}$ (lines cyan and blue) and given $\widetilde{\gamma} = 0.09$,

Second, in Figure 4 we isolate the importance of local knowledge spillovers in R&D $\widetilde{\gamma}$ for aggregate output. To do this, we calculate $\mathbb{E}\left\{\frac{\widehat{Q_d}}{Q_d}\right\}^{\widetilde{\gamma}=\gamma^*} - \mathbb{E}\left\{\frac{\widehat{Q_d}}{Q_d}\right\}^{\widetilde{\gamma}=0}$, where $\mathbb{E}\left\{\frac{\widehat{Q_d}}{Q_d}\right\}^{\widetilde{\gamma}=\gamma^*}$ is the expected change in aggregate output as in Figure 3a, and $\mathbb{E}\left\{\frac{\widehat{Q_d}}{Q_d}\right\}^{\widetilde{\gamma}=0}$ is the expected change in aggregate output in an scenario without local knowledge spillovers in R&D. We observe that the quantitative importance of the spillovers depend on the reduction of inventor migration costs. For $\kappa = \frac{1}{2}$, the spillovers explain between 1% and 3% out of the 22% increase in aggregate output. For $\kappa = 1$, spillovers at baseline value $\widetilde{\gamma} = 0.18$ explain 12.5% out of the 50% increase in aggregate output. This exercise reveals the complementarity when introducing policies that promote local knowledge spillovers in R&D and reduce inventor migration costs to foster economic activity.

Figure 4: Role of $\widetilde{\gamma}$ in reducing inventor migration costs



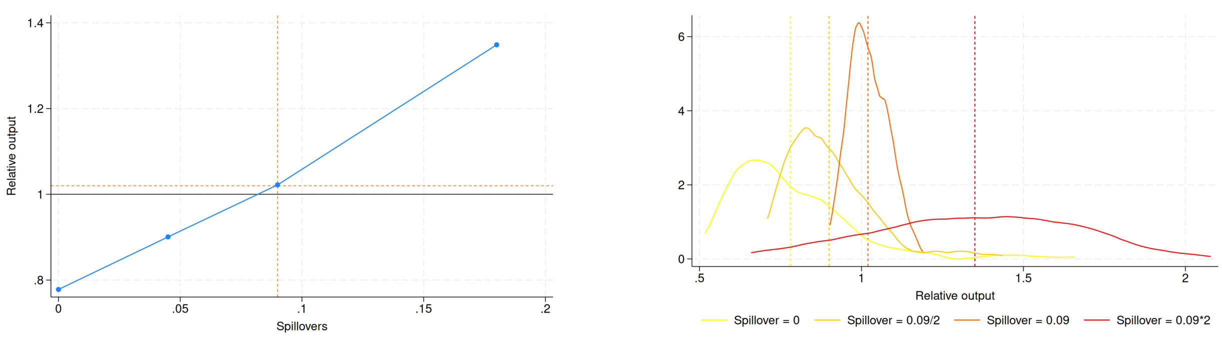
Notes: In Panel 4, the horizontal axis is the proportional reduction parameter of inventor migration costs κ in $\widehat{\mu^R_{od}}$, and the vertical axis is the expected change in aggregate output $\mathbb{E}\left\{\frac{\widehat{Q_d}}{Q_d}\right\}^{\widetilde{\gamma}=\gamma^*} - \mathbb{E}\left\{\frac{\widehat{Q_d}}{Q_d}\right\}^{\widetilde{\gamma}=0}$, where $\mathbb{E}\left\{\frac{\widehat{Q_d}}{Q_d}\right\}^{\widetilde{\gamma}=\gamma^*}$ is the expected change in aggregate output as in Figure 3a, and $\mathbb{E}\left\{\frac{\widehat{Q_d}}{Q_d}\right\}^{\widetilde{\gamma}=0}$ is the expected change in aggregate output in an scenario without local knowledge spillovers in R&D (lines yellow, orange, brown, and red).

4.2 2020 German R&D Tax Allowance Act

In this section, we evaluate the *2020 German R&D Tax Allowance Act*, which introduced a R&D tax incentive scheme as from January 1st 2020. Under this scheme, firms were entitled to receive funding for their R&D activities. In particular, this scheme provides a 25% subsidy for in-house R&D activities regardless of firm characteristics (Deloitte, 2020). In Figure 5, we solve a equilibrium with R&D subsidies as in Section 2.6 and evaluate the effect of implementing a subsidy of $s_o = 25\%, \forall o$ on aggregate output Q_d. In Panel 5a we show that, at a baseline value of spillovers $\widetilde{\gamma} = 0.09$, this policy would have a modest impact of increasing output by 2.2% on average. More importantly, we see that the policy could have increase or reduced output depending on the degree of local knowledge spillovers in R&D. For example, if $\widetilde{\gamma} = 0$ then aggregate output would have decreased by around 21%. In contrast, if $\widetilde{\gamma} = 0.18$ then aggregate output would have increased by around 35%. The mechanism behind this result is because the R&D subsidies are funded by labor taxes. For example, under low or no spillovers, the decrease in income from labor taxes overcomes the benefit from the subsidy, which in turn decreases output in average. In Panel 5b, we study the heterogeneous effects on aggregate output $\frac{\widehat{Q_d}}{Q_d}$ across West German locations for different values of $\widetilde{\gamma}$. We observe that the modest increase in aggregate output mask a decent amount of heterogeneity. For example, at $\widetilde{\gamma} = 0.09$, changes in output across West German locations range from -10% to 25%. This in contrast with other values of $\widetilde{\gamma}$ where the degree of heterogeneity increases with its effect on output.

Figure 5: 25% R&D subsidy

(a) Aggregate effect (b) Heterogenenous effects



Notes: This figure is comprised by two panels. In Panel 5a, the horizontal axis is local knowledge spillovers in R&D $\widetilde{\gamma}$, and the vertical axis is the expected change in aggregate output $\mathbb{E}\left\{\frac{\widehat{Q_d}}{Q_d}\right\}^{\gamma=\gamma^*}$, where $\widehat{Q_d}$ is aggregate output given $s_o = 25\%$, Q_d is aggregate output given $s_o = 0\%$. In Panel 5b we show density plots across West German locations for $\frac{\widehat{Q_d}}{Q_d}$ for spillovers $\widetilde{\gamma} = \{0, \frac{0.09}{2}, 0.09, 0.18\}$ (lines yellow, orange, brown, and red),

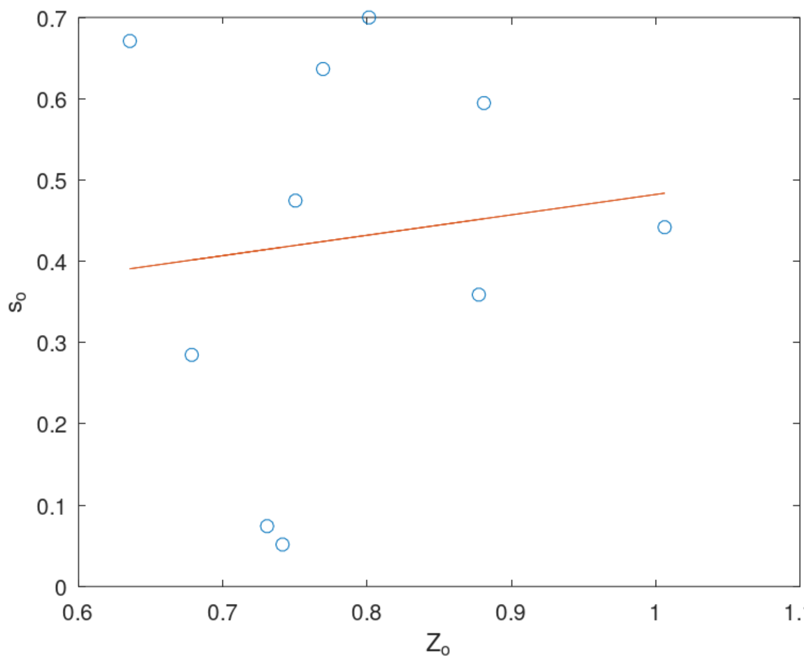
6

4.3 Place-based R&D subsidies

Notice that the 25% subsidy from the *2020 German R&D Tax Allowance Act* is blind to geography. Given the presence of local knowledge spillovers in R&D, it is not obvious that a flat subsidy is optimal or close to optimal. Then, in this section we address whether R&D policies that internalize these spillovers are place-based. To do this, we follow two sequential steps. First, we solve the social planner problem. This yields an optimal allocation $\widehat{X}_o$. Second, we find the set of R&D subsidies s_o^*, which is a least squares problem between the competitive equilibrium with R&D subsidies X_o and the target $\widehat{X}_o$. Details about these steps are in the Supplementary Appendix.

For computational easiness, we consider a simplified economy with 10 locations. In Figure 6, we see that s_o^* vary across locations. More importantly, we see that locations with higher fundamentals for R&D should receive a higher R&D subsidy on average. This highlights the importance of accounting for a country's geography of innovation when implementing R&D subsidies.

Figure 6: Place-based R&D subsidies



Notes: The economy is comprised by 10 locations in West Germany. Horizontal axis are location fundamentals for R&D $\mathcal{Z}_o$. Vertical axis are optimal R&D subsidies s_o^*. Each blue dot is a West German location, and the red line is a linear fit across West German locations.

5 Conclusions

In this paper I quantify the aggregate importance of local knowledge spillovers in R&D. I causally estimate these spillovers by exploiting the historical episode of the arrival of East German inventors across West Germany after the Reunification of Germany. I then embed these spillovers into a spatial model of innovation, and use the model to quantify the importance of these spillovers when implementing policies that promote R&D activities. I show that reducing migration costs for inventors and subsidies to firms' R&D activities can substantially increase aggregate output, and local knowledge spillovers in R&D is crucial for the effectiveness of these policies.

This paper have abstracted from other different channels that could also contribute and interact with these spillovers. First, occupational choice between workers and inventors, or firm selection into R&D through firm heterogeneity could amplify the effect of policies due to entry of agents into innovation. Second, inter-temporal knowledge spillovers could be introduced in the model to quantify the role of local knowledge spillovers in R&D and R&D policies for long-run growth. Finally, new micro-data on inventors also allows to account for the importance of firm-level spillovers and the rise of teams.

References

Abrams, D. S., Akcigit, U., and Grennan, J. (2013). Patent Value and Citations: Creative Destruction or Strategic Disruption? Technical report, National Bureau of Economic Research.

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., and Kerr, W. (2018). Innovation, Reallocation, and Growth. *American Economic Review*, 108(11):3450–91.

Akcigit, U., Hanley, D., and Serrano-Velarde, N. (2021). Back to Basics: Basic Research Spillovers, Innovation Policy, and Growth. *The Review of Economic Studies*, 88(1):1–43.

Allen, T. and Donaldson, D. (2020). Persistence and path dependence in the spatial economy. Technical report, National Bureau of Economic Research.

Arkolakis, C., Lee, S. K., and Peters, M. (2020). European Immigrants and the United States' Rise to the Technological Frontier. Technical report, mimeo.

Arundel, A., van de Paal, G., and Soete, L. (1995). Innovation Strategies of Europe's Largest Industrial Firms. Results of the PACE Survey for Information Sources, Public Research, Protection of Innovations and Government Programme.

Atkeson, A. and Burstein, A. T. (2010). Innovation, Firm Dynamics, and International Trade. *Journal of Political Economy*, 118(3):433–484.

Audretsch, D. B. and Feldman, M. P. (1996). R&D Spillovers and the Geography of Innovation and Production. *The American Economic Review*, 86(3):630–640.

Bahar, D., Choudhury, P., and Rapoport, H. (2020). Migrant Inventors and the Technological Advantage of Nations. *Research Policy*, 49(9):103947.

Black, D. and Henderson, V. (1999). A Theory of Urban Growth. *Journal of Political Economy*, 107(2):252–284.

Blind, K., Edler, J., Frietsch, R., and Schmoch, U. (2006). Motives to Patent: Empirical Evidence from Germany. *Research policy*, 35(5):655–672.

Bloom, N., Schankerman, M., and Van Reenen, J. (2013). Identifying Technology Spillovers and Product Market Rivalry. *Econometrica*, 81(4):1347–1393.

Borjas, G. J. and Doran, K. B. (2012). The Collapse of the Soviet Union and the Productivity of American Mathematicians. *The Quarterly Journal of Economics*, 127(3):1143–1203.

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-Experimental Shift-share Research Designs. *The Review of Economic Studies*, 89(1):181–213.

Bosetti, V., Cattaneo, C., and Verdolini, E. (2015). Migration of Skilled Workers and Innovation: A European Perspective. *Journal of International Economics*, 96(2):311–322.

Broda, C. and Weinstein, D. E. (2006). Globalization and the Gains from Variety. *The Quarterly Journal of Economics*, 121(2):541–585.

Bryan, G. and Morten, M. (2019). The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia. *Journal of Political Economy*, 127(5):2229–2268.

Burchardi, K. B., Chaney, T., Hassan, T. A., Tarquinio, L., and Terry, S. J. (2020). Immigration, Innovation, and Growth. Technical report, National Bureau of Economic Research.

Burchardi, K. B. and Hassan, T. A. (2013). The Economic Impact of Social Ties: Evidence from German Reunification. *The Quarterly Journal of Economics*, 128(3):1219–1271.

Carlino, G. A., Chatterjee, S., and Hunt, R. M. (2007). Urban Density and the Rate of Invention. *Journal of Urban Economics*, 61(3):389–419.

Cohen, W. M., Goto, A., Nagata, A., Nelson, R. R., and Walsh, J. P. (2002). R&D Spillovers, Patents and the Incentives to Innovate in Japan and the United States. *Research policy*, 31(8-9):1349–1367.

Combes, P.-P., Duranton, G., Gobillon, L., and Roux, S. (2010). Estimating Agglomeration Economies with History, Geology, and Worker Effects. In *Agglomeration Economics*, pages 15–66. University of Chicago Press.

Cremers, K., Ernicke, M., Gaessler, F., Harhoff, D., Helmers, C., McDonagh, L., Schliessler, P., and Van Zeebroeck, N. (2017). Patent litigation in Europe. *European Journal of Law and Economics*, 44(1):1–44.

Crews, L. G. (2023). A dynamic spatial knowledge economy.

Davis, D. R. and Dingel, J. I. (2019). A Spatial Knowledge Economy. *American Economic Review*, 109(1):153–70.

Deloitte (2020). Germany: Survey of Global Investment and Innovation Incentives. Technical report, Deloitte.

Desmet, K., Nagy, D. K., and Rossi-Hansberg, E. (2018). The Geography of Development. *Journal of Political Economy*, 126(3):903–983.

Desmet, K. and Rossi-Hansberg, E. (2014). Spatial Development. *American Economic Review*, 104(4):1211–43.

Duguet, E. and Kabla, I. (2000). Appropriation Strategy and the Motivations to Use the Patent System: an Econometric Analysis at the Firm Level in French Manufacturing. In *The Economics and Econometrics of Innovation*, pages 267–305. Springer.

Duranton, G. (2007). Urban Evolutions: The Fast, the Slow, and the Still. *American Economic Review*, 97(1):197–221.

Duranton, G. and Puga, D. (2001). Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products. *American Economic Review*, 91(5):1454–1477.

Duranton, G. and Puga, D. (2019). Urban Growth and its Aggregate Implications. Technical report, National Bureau of Economic Research.

Eaton, J. and Eckstein, Z. (1997). Cities and Growth: Theory and Evidence from France and Japan. *Regional Science and Urban Economics*, 27(4-5):443–474.

Eaton, J. and Kortum, S. (2002). Technology, Geography, and Trade. *Econometrica*, 70(5):1741–1779.

Feldman, M. P. and Kelley, M. R. (2006). The ex ante assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior. *Research policy*, 35(10):1509–1521.

Glaeser, E. L. (1999). Learning in Cities. *Journal of Urban Economics*, 46(2):254–277.

Goolsbee, A. (1998). Does Government R&D Policy Mainly Benefit Scientists and Engineers?

Greenstone, M., Hornbeck, R., and Moretti, E. (2010). Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings. *Journal of Political Economy*, 118(3):536–598.

Griliches, Z. (1991). The Search for R&D spillovers. *National Bureau of Economic Research Working Paper Series*, (w3768).

Gruber, J., Johnson, S., and Moretti, E. (2022). Place-Based Productivity and Costs in Science. Technical report, National Bureau of Economic Research.

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

Hallak, J. C. and Sivadasan, J. (2013). Product and Process Productivity: Implications for Quality Choice and Conditional Exporter Premia. *Journal of International Economics*, 91(1):53–67.

Harhoff, D., Narin, F., Scherer, F. M., and Vopel, K. (1999). Citation Frequency and the Value of Patented Inventions. *Review of Economics and Statistics*, 81(3):511–515.

Hoisl, K., Harhoff, D., Dorner, M., Hinz, T., and Bender, S. (2016). Social Ties or Patent Quality Signals-Evidence from East German Inventor Migration. In *Academy of Management Proceedings*, volume 2016, page 13594. Academy of Management Briarcliff Manor, NY 10510.

Jaffe, A. B., Trajtenberg, M., and Fogarty, M. S. (2000). Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors. *American Economic Review*, 90(2):215–218.

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *the Quarterly Journal of Economics*, 108(3):577–598.

Jaworski, T., Kitchens, C., and Nigai, S. (2020). Highways and globalization. *International Economic Review*.

Kantor, S. and Whalley, A. (2019). Research Proximity and Productivity: Long-term Evidence from Agriculture. *Journal of Political Economy*, 127(2):819–854.

Kelly, M. and Hageman, A. (1999). Marshallian Externalities in Innovation. *Journal of Economic Growth*, 4(1):39–54.

Kerr, W. R. and Kominers, S. D. (2015). Agglomerative Forces and Cluster Shapes. *Review of Economics and Statistics*, 97(4):877–899.

Koehler, M. (2018). Estimating the benefits of r&d subsidies for germany. *ZEW-Centre for European Economic Research Discussion Paper*, (18-002).

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics*, 132(2):665–712.

Koike-Mori, Y., Maruyama, T., and Okumura, K. (2023). Inventor mobility, knowledge diffusion, and growth. *Knowledge Diffusion, and Growth (July 3, 2023)*.

Kortum, S. S. (1997). Research, patenting, and technological change. *Econometrica: Journal of the Econometric Society*, pages 1389–1419.

Kosfeld, R. and Werner, A. (2012). German Labour Markets-New Delineation after the Reforms of German District Boundaries 2007-2011. *RAUMFORSCHUNG UND RAUMORDNUNG-SPATIAL RESEARCH AND PLANNING*, 70(1):49–64.

Krebs, O. and Pflüger, M. (2021). On the road (again): Commuting and local employment elasticities in germany.

Krugman, P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade. *The American Economic Review*, 70(5):950–959.

Lanjouw, J. O. and Schankerman, M. (2004). Patent quality and research productivity: Measuring innovation with multiple indicators. *The economic journal*, 114(495):441–465.

Marshall, A. (1890). Principles of Economics/Royal Economic Society.

Melitz, M. J. (2003). The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity. *Econometrica*, 71(6):1695–1725.

Mestieri, M., Berkes, E., and Gaetani, R. (2021). Cities and Technological Waves.

Monras, J. (2018). Economic Shocks and Internal Migration.

Moretti, E. (2021). The Effect of High-tech Clusters on the Productivity of Top Inventors. *American Economic Review*, 111(10):3328–75.

Moser, P., Voena, A., and Waldinger, F. (2014). German Jewish émigrés and US invention. *American Economic Review*, 104(10):3222–55.

Nagy, D. K. et al. (2016). City Location and Economic Development. *Princeton University, mimeograph.*

Papageorgiadis, N. and Sofka, W. (2020). Patent enforcement across 51 countries–Patent enforcement index 1998–2017. *Journal of World Business*, 55(4):101092.

Pellegrina, H. S. and Sotelo, S. (2021). Migration, Specialization, and Trade: Evidence from Brazil's March to the West. Technical report, National Bureau of Economic Research.

Peri, G., Shih, K., and Sparber, C. (2015). STEM workers, H-1B visas, and productivity in US cities. *Journal of Labor Economics*, 33(S1):S225–S255.

Peters, M. (2022). Market Size and Spatial Growth: Evidence from Germany's Post-war Population Expulsions. *Econometrica*, 90(5):2357–2396.

Pitkethly, R. H. (2001). Intellectual Property Strategy in Japanese and UK Companies: Patent Licensing Decisions and Learning Opportunities. *Research Policy*, 30(3):425–442.

Prato, M. (2021). The Global Race for Talent: Brain Drain, Knowledge Transfer and Growth. *mimeo*.

Ramondo, N., Rodríguez-Clare, A., and Saborío-Rodríguez, M. (2016). Trade, Domestic Frictions, and Scale Effects. *American Economic Review*, 106(10):3159–84.

Redding, S. J. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics*, 101:148–167.

Roca, J. D. L. and Puga, D. (2017). Learning by Working in Big Cities. *The Review of Economic Studies*, 84(1):106–142.

Romer, P. M. (2000). Should the Government Subsidize Supply or Demand in the Market for Scientists and Engineers? *Innovation Policy and the Economy*, 1:221–252.

Schalk, H., Tager, U., and Brander, S. (1999). Wissensverbreitung und Diffusionsdynamik im Spannungsfeld zwischen innovierenden und imitierenden Unternehmen. *Ifo-Institut fur Wirtschafts-forschung, Munchen*.

Silverberg, G. and Verspagen, B. (2007). The size distribution of innovations revisited: an application of extreme value statistics to citation and value measures of patent significance. *Journal of Econometrics*, 139(2):318–339.

Thompson, P. (2006). Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-added Citations. *The Review of Economics and Statistics*, 88(2):383–388.

Trajtenberg, M. (1990). A Penny for your Quotes: Patent Citations and the Value of Innovations. *The RAND Journal of Economics*, pages 172–187.

Trajtenberg, M. (2001). R&D policy in Israel. *Innovation policy in the knowledge-based economy*, pages 409–454.

WEF (2018). The global competitiveness report 2018. World Economic Forum.

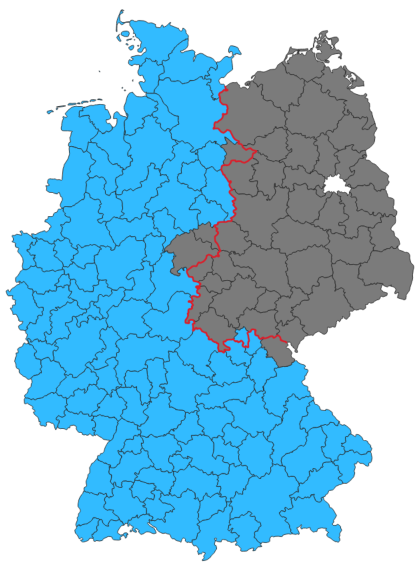
Williams, T. (2023). Right idea, wrong place? knowledge diffusion and spatial misallocation in r&d. *Knowledge Diffusion and Spatial Misallocation in R&D (January 12, 2023)*.

Wilson, D. J. (2009). Beggar Thy Neighbor? The In-state, Out-of-state, and Aggregate Effects of R&D Tax Credits. *The Review of Economics and Statistics*, 91(2):431–436.

Appendix

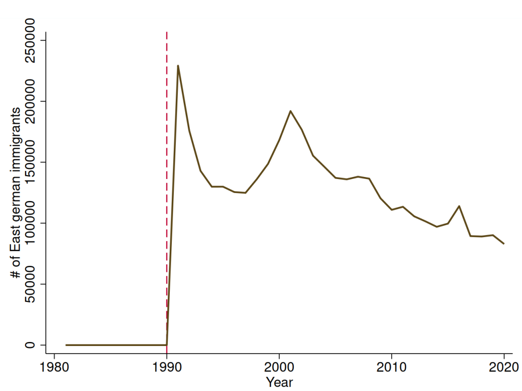
A Additional figures

Figure A.1: West Germany



Notes: This figure shows a map of Germany. West Germany is on the left side of the map (blue), and East Germany is on the right side of the map (gray). The red line separating West and East Germany is the *Iron Curtain*, which was lifted on October 1990. The missing area in East Germany is Berlin. The administrative boundaries are labor markets.

Figure A.2: The Exodus to the West



Notes: This figure shows the yearly number of East Germans migrating to West Germany. The dashed red line denotes 1990, the date of the Reunification of Germany.

B Additional tables

Table A.1: Top 10 West german cities, 2014

	Size		Size
Panel A: Electrical engineering		*Panel D: Mechanical engineering*	
Stuttgart	15.096	Stuttgart	15.865
Munchen	13.776	Munchen	8.140
Regensburg	5.941	Boblingen	5.858
Nurnberg	4.533	Frankfurt	3.577
Erlangen	4.049	Ravensburg	3.318
Karlsruhe	4.005	Erlangen	3.197
Boblingen	2.772	Karlsruhe	3.093
Reutlingen	2.728	Wolsfburg	2.592
Soest	2.552	Dusseldorf	2.540
Frankfurt am Main	2.200	Heilbronn	2.471
Panel B: Instruments		*Panel E: Workers*	
Stuttgart	13.584	Hamburg	6.482
Munchen	8.732	Munchen	5.541
Heidenheim	6.506	Frankfurt	5.369
Erlangen	5.764	Stuttgart	5.070
Boblingen	4.965	Dusseldorf	4.560
Frankfurt	4.109	Koln	3.640
Rottweil	4.052	Essen	3.333
Freiburg	3.424	Hannover	2.541
Regensburg	2.968	Nurnberg	1.932
		Bremen	1.895
Panel C: Chemistry			
Dusseldorf	11.011		
Stuttgart	10.734		
Hamburg	7.202		
Munchen	6.301		
Frankfurt	5.609		
Altotting	2.908		
Essen	2.700		
Koln	2.423		
Reutlingen	2.423		
Erlangen	2.285		

Notes: This table is comprised by five panels. Panels A-D reports the share of inventors working on their corresponding technological area that lives in a given city. Panel E reports the share of workers that lives in a given city. In each panel, I only report the top 10 cities.

Table A.2: OLS models, robustness

	Panel A: $\log(1+Z)$					
	(1)	(2)	(3)	(4)	(5)	(6)
EPO	0.117	0.143	0.224	0.184	0.0859	0.173
	(0.0186)	(0.0173)	(0.0135)	(0.0319)	(0.0349)	(0.0679)
EU	0.142	0.193	0.255	0.203	0.103	0.245
	(0.0208)	(0.0162)	(0.0178)	(0.0461)	(0.0463)	(0.0864)
	Panel B: $IHS(Z)$					
	(1)	(2)	(3)	(4)	(5)	(6)
DPMA	0.0847	0.135	0.118	0.130	0.108	0.217
	(0.0326)	(0.0209)	(0.0205)	(0.0475)	(0.0440)	(0.0798)
EPO	0.140	0.171	0.266	0.223	0.102	0.214
	(0.0219)	(0.0204)	(0.0160)	(0.0389)	(0.0431)	(0.0810)
EU	0.142	0.193	0.255	0.203	0.103	0.245
	(0.0208)	(0.0162)	(0.0178)	(0.0461)	(0.0463)	(0.0864)
$\iota_{d,t}$		✓	✓	✓	✓	✓
$\iota_{a,t}$			✓	✓	✓	✓
ι_{da}				✓	✓	✓
ι_{ω}					✓	✓
ι_{i}						✓
N	177,301	177,300	177,300	177,294	162,803	84,639

Notes: In this table I report OLS estimates from Equation (1). The table is comprised by two panels. In Panel A, the dependent variable is measured as $\log\left(1+Z^{i\omega}_{da,t}\right)$, where $Z^{i\omega}_{da,t}$ is the number of 5-year forward citations. In Panel B, the dependent variable is measured as $IHS\left(Z^{i\omega}_{da,t}\right)$, where $IHS(\cdot)$ is the inverse hyperbolic sine function. Each panel contains a main set of rows denoted by "DPMA", "EPO", and "EU", which indicate the institution that generated the forward citations. The table is comprised by 6 columns. Rows $3, 5, 9, 11, 13$ report the estimate of β, and rows $4, 6, 10, 12, 14$ report standard errors clustered at the (d, a) level. Each column corresponds to a different combination of fixed effects, as pointed out by rows $15-19$. Row 20 report the number of observations.

Table A.3: OLS models, 5-year periods

	(1)	(2)	(3)	(4)	(5)	(6)
$\log(1+Z)$	0.0291	0.0472	0.0449	0.0707	0.0664	0.0907
	(0.0096)	(0.007)	(0.0073)	(0.0146)	(0.0135)	(0.0215)
$IHS(Z)$	0.0368	0.060	0.0568	0.0902	0.0850	0.116
	(0.0124)	(0.0089)	(0.0094)	(0.0187)	(0.0171)	(0.0273)
$\iota_{d,t}$		✓	✓	✓	✓	✓
$\iota_{a,t}$			✓	✓	✓	✓
ι_{da}				✓	✓	✓
ι_{ω}					✓	✓
ι_{i}						✓
N	$177,301$	$177,300$	$177,300$	$177,294$	$162,803$	$84,639$

Notes: In this table I report OLS estimates from Equation (1). Rows $2-3$ report the estimated value of β and its standard errors in parentheses when the dependent variable is measured as $\log\left(1+Z^{i\omega}_{da,t}\right)$, where $Z^{i\omega}_{da,t}$ is the number of 5-year forward citations from the DPMA. Rows $4-5$ report the estimated value of β and its standard errors in parentheses when the dependent variable is measured as $IHS\left(Z^{i\omega}_{da,t}\right)$, where $IHS(\cdot)$ is the inverse hyperbolic sine function. The table is comprised by 6 columns. Each column corresponds to a different combination of fixed effects, as pointed out by rows $6-10$. Standard errors clustered at the (d,a) level. Row 11 reports the number of observations.

Table A.4: IV models, robustness

	Panel A: $\Delta \log(1+Z)$		
	(1)	(2)	(3)
EPO	0.164	0.139	0.209
	(0.0422)	(0.0723)	(0.117)
EU	0.210	0.270	0.343
	(0.0436)	(0.0907)	(0.143)
	Panel B: $\Delta IHS(Z)$		
	(1)	(2)	(3)
DPMA	0.215	0.380	0.498
	(0.0514)	(0.122)	(0.184)
EPO	0.182	0.144	0.237
	(0.0494)	(0.0849)	(0.140)
EU	0.235	0.304	0.393
	(0.0588)	(0.104)	(0.168)
$\iota_{d,t}$		✓	✓
$\iota_{a,t}$			✓
$KP-F$	132.1	34.14	28.23
N	$50,778$	$50,776$	$50,776$

Notes: In this table I report IV estimates from Equation (2). The table is comprised by two panels. In Panel A, the dependent variable is measured as $\Delta \log\left(1+Z^{i\omega}_{da,t}\right)$, where $Z^{i\omega}_{da,t}$ is the number of 5-year forward citations. In Panel B, the dependent variable is measured as $\Delta IHS\left(Z^{i\omega}_{da,t}\right)$, where $IHS(\cdot)$ is the inverse hyperbolic sine function. Each panel contains a main set of rows denoted by "DPMA", "EPO", and "EU", which indicate the institution that generated the forward citations. The table is comprised by 3 columns. Rows $3,5,9,11,13$ report the estimate of β, and rows $4,6,10,12,14$ report standard errors clustered at the (d,a) level. Each column corresponds to a different combination of fixed effects, as pointed out by rows $15-16$. Row 19 shows the first stage Kleibergen-Paap F-statistic (KP-F), and row 20 reports the number of observations.

Table A.5: IV models, 5-year periods

	(1)	(2)	(3)
$\Delta \log(1+Z)$	0.0367	0.0865	0.0849
	(0.0232)	(0.0331)	(0.0428)
$\Delta IHS(Z)$	0.0464	0.109	0.104
	(0.0295)	(0.0420)	(0.0543)
$\iota_{d,t}$		✓	✓
$\iota_{a,t}$			✓
$KP-F$	85.96	26.64	38.15
N	$100,234$	$100,228$	$100,228$

Notes: In this table I report IV estimates from Equation (2). Rows $2-3$ report the estimated value of β and its standard errors in parentheses when the dependent variable is measured as $\Delta \log\left(1+Z^{i\omega}_{da,t}\right)$, where $Z^{i\omega}_{da,t}$ is the number of 5-year forward citations from the DPMA. Rows $4-5$ report the estimated value of β and its standard errors in parentheses when the dependent variable is measured as $\Delta IHS\left(Z^{i\omega}_{da,t}\right)$, where $IHS(\cdot)$ is the inverse hyperbolic sine function. The table is comprised by 3 columns. Each column corresponds to a different combination of fixed effects, as pointed out by rows $6-7$. Standard errors clustered at the (d,a) level. Row 8 shows the first stage Kleibergen-Paap F-statistic (KP-F), and row 9 reports the number of observations.

Table A.6: Estimation of migration costs, 5-year periods

	$n=R$		$n=L$	
	OLS	PPML	OLS	PPML
$\log(dist_{od})$	-1.020	-1.381	-1.505	-1.380
	(0.010)	(0.017)	(0.025)	(0.015)
ρ_1^n	0.481	0.651	0.709	0.650
R^2	0.826	·	0.835	·
N	$17,283$	$54,080$	$43,835$	$54,080$

Notes: In this table I report migration cost elasticities from Equation (34) under 5-year periods. Columns $2-3$ are the regressions for inventors, where column 2 are OLS estimates, and column 3 are PPML estimates. Columns $4-5$ are the regressions for workers, where column 4 are OLS estimates, and column 5 are PPML estimates. For OLS estimates, the dependent variable is measured as $\log\left(\eta^n_{od,t}\right)$ is the log of the share of inventors or workers from o that moved to d during a given period, where I consider 5-year periods. Row 3 is the estimate associated to $\log(dist_{od})$, where $dist_{od}$ is the Euclidean distance in miles from o to d. Row 4 are standard errors two-way clustered at the o,t and d,t level. Row 5 is the implied migration elasticity from the estimates from row 3 given $\kappa = 2.12$. Rows $6-7$ contain the goodness of fit and number of observations in each specification, respectively.