Spatial Knowledge Spillovers in R&D and Aggregate Productivity: Evidence from the Reunification of Germany*

Brian Cevallos Fujiy†
University of Michigan

JOB MARKET PAPER

[Click here for the latest version of the paper]

Abstract

I quantify the importance of spatial knowledge spillovers in Research and Development (R&D) for aggregate productivity. Using new administrative dataset from German inventors, I causally 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. Increasing the number of inventors in a location by 10% leads to average inventor productivity gains of around 4.09%. I then embed these estimated spillovers into a quantitative spatial model of innovation, and use it to quantify the productivity gains from implementing policies that promote R&D activities. First, reducing inventor migration costs by 25% increases aggregate productivity by 5.87%. Second, 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 spatial knowledge spillovers in R&D.

Keywords: Aggregate productivity, agglomeration, spillovers, innovation

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

*Preliminary and incomplete. 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 Dominick Bartelme, Jaedo Choi, Nadim Elayan, Gaurav Khanna, Ryan Kim, Jose Moran, Dev Patel, Heitor Pellegrina, and Linda Tesar. Thanks to participants at UM's International-Macro Lunch Workshop. 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.

†Email: bcfujiy@umich.edu | 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 world-wide 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 were 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 core explanation for why economic activity concentrates. Nevertheless, the extent, causes and consequences of *spatial 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) are there spatial knowledge spillovers in R&D? (ii) are they quantitatively important for aggregate productivity? and (iii) how effective are policies that foster R&D activities to increase aggregate productivity under these spillovers?

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 spatial 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 spatial 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 spatial knowledge spillovers in R&D are large and that they are important for aggregate productivity. First, using new data on German inventors, I causally estimate such spillovers. I perform this task 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 model of innovation that account for the spatial 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 spatial 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 the number of inventors working in a given technological cluster—cluster size, 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 spatial 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 estimate is robust to different specifications of inventor productivity and time aggregation.

I then address potential endogeneity concerns that potentially introduce biases when estimating spatial knowledge spillovers in R&D. On one side, if inventors start working in a certain technological area at a given period due to unobservable idiosyncrasies, so they are initially less productive, and these inventors move to large clusters due to better career prospects, this would introduce a downward bias. On the other side, if there are unobservable cluster-level shocks that increase both cluster size due to immigration of inventors and the productivity of inventors in that cluster, this would introduce an upward bias. Finally, measurement error could introduce a downward bias.

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 cluster size in West Germany, which I then use to causally estimate spatial knowledge spillovers in R&D. The "shifts" are leave-out shocks that measure the total number of inventors that moved from each location

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

in East Germany towards West German clusters, except for the instrumented cluster. The identification assumption is that these shocks are as-good-as-randomly assigned (Borusyak et al., 2022). These shocks are then weighted by exposure "shares" that help predicting how many inventors will arrive to each West German cluster. These shares are constructed based on 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%.

In the second part of the paper, 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 to produce the inputs and hire inventors that engage in R&D to determine the quality of the input. Following Kortum (1997), I model inventors' R&D as the process where they produce ideas that are heterogeneous in productivity. Moreover, their R&D process is subject to the spatial knowledge spillovers I estimated in the data, so inventors that work in locations with a higher density of inventors produce more productive ideas. Workers and inventors are mobile across locations, so they choose where to work according to real wages, amenities, and migration costs. Finally, to avoid that all inventors agglomerate in a single location, I introduce firm-level decreasing returns to R&D which generates local congestion in R&D.

The main prediction of the model is that a location's productivity is endogenously determined by three forces. First, locations with better production fundamentals are more productive. More importantly, locations that hold more inventors are more productive due to spatial knowledge spillovers in R&D. Second, locations that exhibit higher labor costs are less productive since firms are less able to hire them 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 spatial knowledge spillovers in R&D for aggregate productivity. First, I 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 spatial 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 patents a firm filed on 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 spatial 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 spatial 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 spatial knowledge spillovers for inventors. I contribute to this literature by exploiting a historical natural experiment to causally estimate spatial 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). I contribute to this literature by building a quantitative framework that explicitly accounts for spatial knowledge spillovers in R&D I estimate in the data.

Third, this paper contributes to the literature on policies that promote productivity and economic growth. This paper focuses on policies that foster labor mobility and R&D activities. These literatures show that labor mobility matters for productivity and economic growth 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; Pellegrina and Sotelo, 2021; Prato, 2021), and that R&D policies can promote productivity (Goolsbee, 1998; Romer, 2000; Wilson, 2009; Acemoglu et al., 2018; Akcigit et al., 2021). I contribute to this literature by providing a quantitative framework to quantify the productivity gains of implementing migration and R&D policies in general equilibrium.

The remainder of this paper is structured as follows. Section 1 explains how I estimate spatial 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 Spatial Knowledge Spillovers in R&D

In this section I describe the estimation of spatial knowledge spillovers in R&D. The first part of this section describes the data, the second part explains the estimation strategy, and the final part discusses assumptions and results thought this section. Appendices A-C contain additional tables and figures, and details about the data sources.

1.1 Data sources

Linked Inventor Biography (INV-BIO). The key dataset is the INV-BIO dataset 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 $150,000$ inventors in Germany 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. For more details of the data, see Appendix C.

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 spatial knowledge spillovers in R&D, I stack the data in 10-year periods.

West German clusters. I define a 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 $104 \times 5 = 520 \ (d, a)$ 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. This is because knowledge spillovers with other inventors would 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 1985 and 1994, 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 a patent 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 an inventor's 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 and patent an inventor filed was 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 spatial knowledge spillovers in R&D, I construct two main variables. First, I measure inventor productivity $Z^{i\omega}_{da,t}$ at 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 that work in cluster (d, a) at the end of period t.

Additional variables. I construct four additional variables I use for both estimation of spatial knowledge spillovers in R&D in Section 1.3 and calibration exercises 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 the number of filed patents in technological area a such that $\sum_a TechComp_{da} = 1, \forall d$. Third, I measure migration flows during a given period between every location pair $\left\{\nu^L_{od,t}, \nu^R_{od,t}\right\}$ 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 spatial knowledge spillovers, I consider a 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 spatial 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-level shocks that drive the overall productivity of a location. $\iota_{a,t}$ are technological area-period fixed effects that control for overall technological shocks. ι_{da} are cluster fixed effects that account for time-invariant cluster size, and for the fact that some clusters file more patents than others in average. ι_{ω} are establishment fixed effects that account for productivity arising from inventors working for more productive establishments than others. ι_i are inventor fixed effects that control for spatial 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 $(1) - (5)$ show the value of the estimated spillovers as I progressively include the aforementioned fixed effects. The value of these estimates remain around 10%. 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 increased by 10% or moved to a cluster with 10% as many 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 B.2 contains the estimated spillovers under different specifications of inventor productivity. Since the specification with all the fixed effects is the most important one, I focus the discussion on 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 17.3% and 24.5%, 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 inventors's total amount of citations. Column (6) shows that the elasticity of inventor productivity to cluster size is 21.7%. Additionally, when patent citations arose from the EPO and the EU, elasticities are around $21\% - 24\%$.

Finally, results also hold under shorter time horizons. In Table B.3 I show the estimated spillovers when the frequency of the data is 5-year periods, where the first row measures inventor productivity as $log\,(1+x)$, and the third row measures it as $IHS\,(x)$. In both cases, column (6) shows that the spillovers are around 10%. In general, and in line with Moretti (2021), estimated spatial knowledge spillovers decrease with time frequency.

1.3.2 IV approach

Endogeneity concerns. To causally estimate β from Equation (1), the key identification assumption is that the unobservables $\epsilon^{i\omega}_{da,t}$ are uncorrelated to cluster size $R_{da,t}$. Nevertheless, there are at least two concerns that could potentially violate this assumption. First, unobserved time-varying idiosyncratic shocks can bias the estimate of β. Inventors can decide to start working in a certain technological area at a given period due to unobservable reasons, so they initially file a low amount of patents, or file low-quality patents. If these inventors decide to work in large clusters due to better career prospects, this would introduce a downward 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 inventors moving to work in 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.

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 inventors from East Germany 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 between 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 so the process for reunifying 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 the ideal experiment to causally estimate spatial 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 will be sufficient to causally estimate spatial knowledge spillovers in R&D.

First, I use variation in the arrival of inventors towards West German clusters, so the second stage regression is the first-difference of Equation (1):

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

Notice that the only fixed effects in Equation (2) that prevail after introducing first-differences are location-period $\iota_{d,t}$ and technological area-period $\iota_{a,t}$. Notice that $\iota_{d,t}$ is crucial to account for the overall migration towards West German locations. 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, 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 $g_{o,t} \equiv \log\left(\Delta R_{o,t}^{-d,-a}\right)$ is the log of the number of inventors in o that moved to some West German cluster except the instrumented cluster (d,a) during period t. 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.

Following Borusyak et al. (2022), the identification assumption for estimating β is that the overall arrival of East German inventors in West Germany $g_{o,t}$ excluding the instrumented cluster is uncorrelated to unobservables shocks $\epsilon^{i\omega}_{da,t}$. Since these shocks do not include the instrumented cluster (d, a), it is safe to assume that these overall shifts are not correlated to local unobservable demand shocks or unobserved idiosyncratic inventor patenting motives.

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 help predicting which technological area an East German inventor will work on when arriving to a West German cluster. The shares are constructed and then normalized such that $\sum_{o\in\mathcal{E}} s_{o,da} = 1, \forall d, a$.

Table 2 contains the IV estimates of spatial knowledge spillovers in R&D. All the estimates exhibit an F-statistic well-above 10, which reflects the relevance of the instrument. Column (1) reports the estimate of the spillovers when I do not consider any fixed effects. This reports a value of 17.8% 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 30.9%. Finally, it is also key to include technological area-period fixed effects to control for industrial changes after the Reunification. Column (3) contains the main empirical result of this paper: an inventor whose cluster increased by 10% or moved to a cluster with 10% as many inventors obtains productivity gains of 4.09% in average. This estimate is between 2 and 3 times the OLS estimate of 17.5% from column (6) in Table (1). This 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 B.4 contains the estimated spillovers under different specifications of inventor productivity. Since the specification with both location-period and technological area-period fixed effects is the main one, I focus the discussion on 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 20.9% and 34.3%, 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 Inverse Hyperbolic Sine (IHS) for inventors's total amount of citations. Column (3) shows that the elasticity of inventor productivity to cluster size is 49.8%. Additionally, when patent citations arose from the EPO and the EU, elasticities are around $23\%-39\%$.

Finally, results also hold under shorter time horizons. In Table B.5 I show the estimated spillovers when the frequency of the data is 5-year periods, where the first row measures inventor productivity as $\Delta log\left(1+x\right)$, and the third row measures it as $\Delta IHS\left(x\right)$. In both cases, column (3) shows that the spillovers are around 9%.

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 B.2 and B.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 spatial 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 20% higher in a US metropolitan area with twice the population density of any other metropolitan area. My baseline estimate of 40.9% is substantially 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 spatial 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 6.7%, while my estimate from Table 1 is 17.5%. When running the model in first differences, his IV estimate is around 4.9%, while my estimates from Table 2 is 40.9%. Even thought both of these papers estimate spatial knowledge spillovers in R&D at the inventor level, 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 spatial knowledge spillovers in R&D in Germany in comparison to the US.

2 Model

In this section I build a quantitative model of innovation with spatial knowledge spillovers in R&D. Appendices D-E contain details about the derivations in 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) a mass of intermediate input firms. 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). Each input is produced by a single firm, it is tradable across locations, it is produced by workers, and its quality is determined through R&D by inventors. Every input is exported to every location, and I assume a unit mass of intermediates in each location.

Agents. There are two types of agents in each location: (i) workers, and (ii) inventors. Each agent supplies labor inelastically, earns income from wages and redistributed profits from firms, consumes the local final good, and it is mobile across locations. Since agents are mobile, they solve a location choice problem, where each agent living in o decides which d to move to by maximizing their utility.

2.2 Technology

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

$$Q_d = \left(\sum_o \int_{\omega \in \Omega_{od}} A_o^{\omega \frac{1}{\sigma}} Q_{od}^{\omega \frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \tag{4}$$

where Ω_{od} is the set of intermediate input firms in o selling to d, Q_d is the production of the final good, Q_{od}^{ω} is the quantity of intermediate input ω, A_o^{ω} is input quality, and $\sigma > 1$ is the CES across intermediate inputs. The final good firm maximizes profits subject to Equation (4), which yields the demand for intermediate inputs

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

where $P_d^{1-\sigma} = \sum_o P_{od}^{1-\sigma}$ and $P_{od}^{1-\sigma} = \int_{\omega \in \Omega_{od}} A_o^{\omega} P_{od}^{\omega^{1-\sigma}} d\omega$ are CES price indices, and $X_d = P_d Q_d$ is total expenditure on the final good in d.

Intermediate input firms. In each location o, there is a mass of firms that produce a unique and tradable intermediate input. The profits of firm ω selling to d is

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

where w_o^L are worker wages, L_{od}^{ω} is labor demand by ω, and $\tau_{od} > 1$ are iceberg trade costs. A unit of labor is required to produce an intermediate input:

$$L_{od}^{\omega} = Q_{od}^{\omega}. \tag{7}$$

Then, firm ω maximizes total profits subject to Equations (5), (6), and (7):

$$\begin{aligned}\max_{\left\{P^{\omega}_{od},Q^{\omega}_{od},L^{\omega}_{od}\right\}} \pi^{\omega}_{o} &= \sum_{d} \pi^{\omega}_{od} \\ s.t.& \\ \pi_{od} &= P^{\omega}_{od}Q^{\omega}_{od} - \tau_{od} w^{L}_{o} L^{\omega}_{od}, \\ L^{\omega}_{od} &= Q^{\omega}_{od}, \\ Q^{\omega}_{od} &= A^{\omega}_{o} P^{\omega^{-\sigma}}_{od} P^{\sigma-1}_{d} X_{d}.\end{aligned} \tag{8}$$

Then, firms charge a constant markup:

$$P^{\omega}_{od} = \overline{m}\tau_{od} w^{L}_{o}, \forall \omega \in \Omega_{od} \tag{9}$$

where $\overline{m} \equiv \frac{\sigma}{\sigma-1}$ is the CES constant markup over marginal costs. Plugging back Equation (9) in (8), firm total profits are

$$\pi^{\omega}_{o} = \frac{1}{\sigma} A^{\omega}_{o} \sum_{d} \left(\frac{P^{\omega}_{od}}{P_{d}}\right)^{1-\sigma} X_{d}. \tag{10}$$

From (10), we notice that total profits of firm ω increase with market demand from every location d and the quality of its intermediate input A^{ω}_{o}. This is because inputs of higher quality exhibit higher demand from every final good firm.

Quality of intermediate inputs. Each firm ω in every location owns a blueprint that describes the production process of intermediate input with quality A^{ω}_{o}. The blueprint is comprised by n^{ω}_{o} ideas generated by firm's inventors, and ideas are heterogeneous in productivity. Then, the quality of the intermediate input is

$$A^{\omega}_{o} = \mathbb{Z}^{\omega}_{o} n^{\omega}_{o}, \tag{11}$$

where $\mathbb{Z}^{\omega}_{o}$ is the expected productivity of inventors' ideas. In Appendix D.2 I provide two microfoundations that generate isomorphic expressions for the quality of intermediates. To provide intuition on why the quality of the intermediate input is determined by the expected productivity of inventors' ideas, I sketch the first microfoundation based on complementary tasks.

The intermediate input ω is produced by following continuum of necessary tasks contained within its blueprint. Firm's inventors generate n^{ω}_{o} ideas that determine the quality of each task within the ω's blueprint. Since tasks are necessary, ideas are heterogeneous in productivity, and each idea improves all tasks' quality, then the quality of the intermediate is determined by the expected productivity of ideas. Finally, the firm hires R_o^ω inventors who in turn generate $n_o^\omega \leq R_o^\omega$ ideas. That is, I consider decreasing returns to R&D, such that

$$n_o^\omega = R_o^{\omega^\zeta}, \tag{12}$$

where $\zeta \in (0,1)$ is the degree of decreasing returns to R&D. This assumption introduces a local congestion in R&D, which is key to countervail agglomeration forces in R&D, and therefore avoid the possibility of all inventors concentrating in a single location.

Productivity of ideas. Inventors produce ideas according to the innovation process from Kortum (1997). Each inventor i hired by firm ω generates an idea to be implemented in the blueprint of the intermediate input. Ideas are heterogeneous in productivity $Z_o^{i\omega}$ drawn from a probability distribution:

$$Z_o^{i\omega} \sim Frechet\left(\alpha, \lambda_o^{\frac{1}{\alpha}}\right), \tag{13}$$

where α and λ_o are the shape and scale parameters of the Frechet distribution, respectively. Appendix D.1 describes the inventors' innovation process that generate a Frechet distribution for the productivity of ideas. λ_o is the *spillover function* since it embeds economic forces that are exogenous to inventors and increase their productivity in average. Guided by the empirical evidence on spatial knowledge spillovers in R&D in the previous section, I consider the following functional form:

$$\lambda_o^{\frac{1}{\alpha}} = \mathcal{A}_o R_o^{\widetilde{\gamma}}, \tag{14}$$

where $\mathcal{A}_o$ is a fundamental location productivity, R_o is the total number of inventors in o (i.e. *cluster size*), and $\tilde{\gamma} \equiv \frac{\gamma}{\alpha}$ are spatial knowledge spillovers in R&D.[2] Finally, considering the probability distribution in Equation (13), then the expected productivity of inventors' ideas is

$$\mathbb{Z}_o^\omega = \psi \lambda_o^{\frac{1}{\alpha}}, \tag{15}$$

where $\psi > 0$ is a constant that arises from the microfoundation for the quality of intermediate inputs.

Research and Development (R&D). Each firm ω in every location engages in R&D, where each firm optimally decides how many inventors to hire. The optimal number of inventors arises from the trade-off between the cost of hiring inventors and higher quality. Then, firm

[2]Technically, γ are spatial knowledge spillovers in R&D. Since this parameter is not separable from the shape parameter α, I consider $\tilde{\gamma} \equiv \frac{\gamma}{\alpha}$ throughout the paper.

ω maximizes total profits after R&D expenditure subject to (10), (14), (11), and (15):

$$\max_{\{R_o^\omega\}} \overline{\pi}_o^\omega = \pi_o^\omega - w_o^R R_o^\omega \tag{16}$$

$$s.t.$$

$$\pi_o^\omega = \frac{1}{\sigma} A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d,$$

$$A_o^\omega = \psi \mathcal{A}_o R_o^{\widetilde{\gamma}} R_o^{\omega\zeta}.$$

Then, firms' demand for inventors is

$$R_o^\omega = \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}. \tag{17}$$

2.3 Location choice

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

$$U_d^n = \frac{\mathcal{B}_d^n w_d^n \left(1 + \overline{\pi}\right)}{P_d} \quad , n = \{L, R\}, \tag{18}$$

where $\mathcal{B}_d^n$ are type-specific location amenities, and $\overline{\pi}$ are redistributed profits per-capita. There is an exogenous allocation of workers and inventors across locations $\left\{\overline{L}_o, \overline{R}_o\right\}_{\forall o \in \mathcal{S}}$. Then, each agent i of type n working in o moves to d by maximizing its *ex ante* indirect utility:

$$U_{od}^{i,n} = \max_{d \in \mathcal{S}} \left\{\frac{U_d^n}{\mu_{od}^n} \times \epsilon^i\right\} \quad , n = \{L, R\}, \tag{19}$$

where $\mu_{od}^n > 1$ are *iceberg* migration costs, $G(\epsilon) = \exp\left(-\epsilon^{-\kappa}\right)$ are location preference shocks, and κ is the migration elasticity. Given Equations (18)-(19), the share of agents of type n moving from o to d is

$$\eta_{od}^n = \frac{\left(\frac{U_d^n}{\mu_{od}^n}\right)^\kappa}{\sum_\delta \left(\frac{U_\delta^n}{\mu_{o\delta}^L}\right)^\kappa} \quad , n = \{L, R\}. \tag{20}$$

2.4 Aggregate variables

Aggregate productivity. In the model, a location's productivity is defined as the average quality of intermediates in a location. Since firms are symmetric, then from Equations (14)-

(12) and (17), location's productivity is

$$A_o^{1-\zeta} \propto \underbrace{\left(\mathcal{A}_o R_o^{\tilde{\gamma}}\right)}_{spillovers} \underbrace{\left(w_o^{L^{\sigma-1}} w_o^R\right)^{-\zeta}}_{labor\ costs} \underbrace{\left(\sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d\right)^{\zeta}}_{market\ access}. \tag{21}$$

Equation (21) is the main expression of the model because it encapsulates the market forces that endogenously determine aggregate productivity. First, locations that exhibit higher fundamentals for productivity $\mathcal{A}_o$, exhibit higher R&D employment R_o are more productive. This is a direct implication of firms' inventors' R&D being subject to spatial knowledge spillovers. Second, locations that report higher labor costs $\left\{w_o^L, w_o^R\right\}$ are less productive. This is the result of firms' trade-off for both the production of their inputs and R&D. On one side, firms want to hire inventors and workers to increase the quality of their inputs and sell them to all locations. On the other, higher labor demand increases overall labor costs, which makes hiring more difficult for firms. Third, locations with higher market access are more productive. This term arises because intermediate inputs are tradable. Intuitively, higher demand for inputs pushes firms' incentives to increase the quality of their inputs. Finally, I define aggregate productivity as the average quality of intermediate inputs in all the economy.

Price indices. Given Equations (9) and (21), price indices are

$$P_{od}^{1-\sigma} = A_o \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma} \quad and \quad P_d^{1-\sigma} = \sum_o A_o \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma}. \tag{22}$$

Trade shares. Given Equations (22), location o's share in location d's expenditure is

$$\chi_{od} = \frac{A_o \left(\tau_{od} w_o^L\right)^{1-\sigma}}{\sum_o A_o \left(\tau_{od} w_o^L\right)^{1-\sigma}}. \tag{23}$$

Equation (23) shows that trade shares depend directly on locations' productivity A_o; that is, a higher A_o increases location o's share in d's total expenditure. Intuitively, the higher quality of intermediate inputs from o increases their demand from all other locations, which in turn increases o's trade share. Moreover, some of the concentration of economic activity is explained by agglomeration forces arising from locations' R&D activity.

Profits per-capita. Firms' profits are invested in a national fund, and they are then redistributed uniformly across the country's population. Then, plugging Equation (17) back in firm profits (16), yields location's total profits:

$$\overline{\pi}_o = \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma}\psi}{\sigma} \frac{\mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R\zeta} w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}. \tag{24}$$

Then, profits per-capita are

$$\overline{\pi} = \frac{1}{N} \sum_o \overline{\pi}_o. \tag{25}$$

2.5 Equilibrium

Inventors market. From firm's demand for inventors (17), location o's aggregate demand for inventors is

$$w_o^R = \frac{\zeta \psi \overline{m}^{1-\sigma}}{\sigma} \frac{\mathcal{A}_o R_o^{\widetilde{\gamma}-(1-\zeta)}}{w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d. \tag{26}$$

Notice in Equation (26) that, if $\widetilde{\gamma}$ is sufficiently high, then the demand for inventors is upward sloping. That is, if spatial knowledge spillovers in R&D are higher than the decreasing returns to R&D, then the demand for inventors will exhibit an upward slope. This is a similar mechanism as in Allen and Donaldson (2020); Krugman (1979), where a sufficiently strong productivity spillovers can lead to non-unique equilibria. I later show that the calibration of the model rules out the possibility of multiple equilibria. Now, given Equation (20), the supply of inventors in each location is

$$R_d = \sum_o \eta_{od}^R \overline{R}_o. \tag{27}$$

Workers market. From Equations (5), (7), and (9), the aggregate demand for workers is

$$w_o^L = \overline{m}^{-1} \left(\frac{A_o}{L_o} \sum_d \tau_{od}^{-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{\sigma}}. \tag{28}$$

Given Equation (20), the equilibrium number of workers and inventors in each location is

$$L_d = \sum_o \eta_{od}^L \overline{L}_o. \tag{29}$$

Equilibrium in goods market. To close the model, in every location, total income equals total expenditure. Income Y_o is comprised by wages earned by workers and inventors:

$$Y_o = (1 + \overline{\pi}) \left(w_o^L L_o + w_o^R R_o \right). \tag{30}$$

Expenditure X_o is comprised by purchased intermediates from every location d:

$$X_o = \sum_d \chi_{od} X_d. \tag{31}$$

In equilibrium, given Equation (23), income equals expenditure $X_o = Y_o, \forall o$:

$$w_o^L L_o + w_o^R R_o = \sum_d \chi_{od} \left(w_d^L L_d + w_d^R R_d \right) \tag{32}$$

Government budget. A national government implements a set of subsidies s_o that are funded with a uniform labor tax τ. I assume a balanced government budget constant, such that

$$\tau \sum_o \left(w_o^L L_o + w_o^R R_o \right) = \sum_o s_o \left(w_o^R R_o \right). \tag{33}$$

Definition 1 (Equilibrium). *Given iceberg trade costs* $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$*, iceberg migration costs* $\{\mu_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$*, location fundamentals* $\{\mathcal{A}_o, \mathcal{B}_o^n\}_{\forall o \in \mathcal{S}}^{n=\{L,R\}}$*, an equilibrium is a set of wages* $\{w_o^n\}_{\forall o \in \mathcal{S}}^{n=\{L,R\}}$*, prices* $\{P_o\}_{\forall o \in \mathcal{S}}$*, quantities* $\{L_o, R_o, Q_o\}_{\forall o \in \mathcal{S}}$*, and aggregate productivity* $\{A_o\}_{\forall o \in \mathcal{S}}$ *such that (i) inventor markets clear (Equations (26) and (27)), (ii) worker markets clear (Equations (28) and (29)), (iii) goods market clear (Equation (32)), and (iv) locations' productivity are determined by (21).*

2.6 Equilibrium with R&D subsidies

To evaluate the implementation of R&D subsidies, I consider a set of subsidies s_o that are funded with a uniform labor tax τ. First, firms' total profits after R&D is $\overline{\pi}_o^\omega = \pi_o^\omega - (1 - s_o) w_o^R R_o^\omega$, so the subsidy acts as a positive shock for the demand of inventors which value. Second, since these subsidies are funded through a uniform labor tax, then workers and inventors' income are now $w_o^n (1 + \overline{\pi} - \tau)$. Finally, I assume a balanced government budget constant, so government's income $\tau \sum_o \left(w_o^L L_o + w_o^R R_o \right)$ equals its expenditure $\sum_o s_o \left(w_o^R R_o \right)$.

3 Taking the Model to the Data

In this section I describe the calibration strategy of the model. The model is parametrized by spatial knowledge spillovers in R&D ($\widetilde{\gamma}$), decreasing returns to R&D (ζ), migration costs $\{\mu_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$, location fundamentals $\{\mathcal{A}_o, \mathcal{B}_o^n\}_{\forall o \in \mathcal{S}}^{n=\{L,R\}}$, trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, and a set of additional parameters $\{\alpha, \kappa, \sigma\}$. Table 5 at the end of this section summarizes the calibration strategy of the model, and further details on the parametrization are in Appendix F.

Spatial knowledge spillovers in R&D $\{\widetilde{\gamma}\}$. The reduced-form estimates for spatial knowledge spillovers in R&D in Section 1 can be mapped directly to $\widetilde{\gamma}$. Equation (15) describes a direct relationship between the expected productivity of inventors and the spillovers they are subject to. Considering the spatial knowledge spillovers from Equation (14), and after introducing logs and first-differences, the model yields a log-log relationship between inventor productivity and cluster size:

$$\Delta \log \left(Z_o^{i\omega}\right) = \widetilde{\gamma} \Delta \log \left(R_o\right) + \Delta \epsilon_o^{i\omega}. \tag{34}$$

Then, to create a direct mapping between $\widetilde{\gamma}$ in Equation (34) and the estimated spillovers $\beta = 0.409$ in Equation (2), is necessary an elasticity between productivity and number of forward citations. Following Lanjouw and Schankerman (2004), I consider an elasticity of 0.22, so I set $\widetilde{\gamma} = (0.22)\,\beta \approx 0.09$.

Decreasing returns to R&D $\{\zeta\}$. I express Equation (12) in logs:

$$\log \left(n_o^{\omega}\right) = \zeta \log \left(R_o^{\omega}\right) + \epsilon_o^{\omega}, \tag{35}$$

where ϵ_o^{ω} are i.i.d shocks. Then, using data on the number of firms' inventors R_o^{ω} and the number of firms' inventors that generated an idea n_o^{ω}, ζ is can be estimated directly from the data. To keep the estimation consistent with the reduced-form estimates, I consider 10-year periods. To account for additional confounding factors, in Equation (35) I also include establishment fixed effects ι_{ω}, and location-period fixed effects ι_o. Column (3) of Table 3 shows the main estimate of $\zeta = 0.65$, which confirms the existence of decreasing returns to R&D. This estimate is within the upper bound of previous values in the literature between 0.1 and 0.6 (Kortum, 1993). This is because the time frequency of the estimate is long in comparison to previous estimates that rely on either yearly or cross-sectional variation. Additional estimations are in Table B.7, which also exhibit decreasing returns to R&D.

Table 3: Estimation of decreasing returns to R&D

	(1)	(2)	(3)
$\log\left(R_o^\omega\right)$	0.718	0.704	0.65
	(0.009)	(0.0096)	(0.0103)
ι_o		✓	✓
ι_ω			✓
N	$49,297$	$49,297$	$25,010$
R^2	0.72	0.812	0.904

Notes: In this table I report estimates for decreasing returns to R&D from Equation (35). To account for zeros on the dependent variable, I consider $\log\left(1+n_o^\omega\right)$. Each column is an specification with different combinations of fixed effects. The fixed effects included in each specification are determined by rows $4-5$. Row 2 contains the estimates for ζ, and row 3 contain standard errors, which are clustered at the ω level. Rows $6-7$ contain the number of observations and goodness of fit in each specification, respectively.

Migration costs $\{\mu_{od}^n\}$. For each agent type $n=\{L,R\}$, 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 $\left\{\rho_0^L,\rho_0^R\right\}=\{1.361,1.354\}$. To estimate $\{\rho_1^n\}$, the location choice problem of the model yields migration gravity equations for both workers and inventors:

$$\log\left(\eta_{od}^n\right)=\iota+\iota_o+\iota_d-\kappa\rho_1^n\log\left(dist_{od}\right)+\epsilon_{od}^n, n=\{L,R\}\,. \tag{36}$$

The gravity equation in (36) states that, conditional on origin and destination fixed effects $\{\iota_o,\iota_d\}$, data on geographic distance between locations, and the migration elasticity κ, the 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 4, I consider $\left\{\rho_1^L,\rho_1^R\right\}=\{0.5915,0.6023\}$. Additional estimates in Table B.6 also confirm the presence of migration costs for workers and inventors.

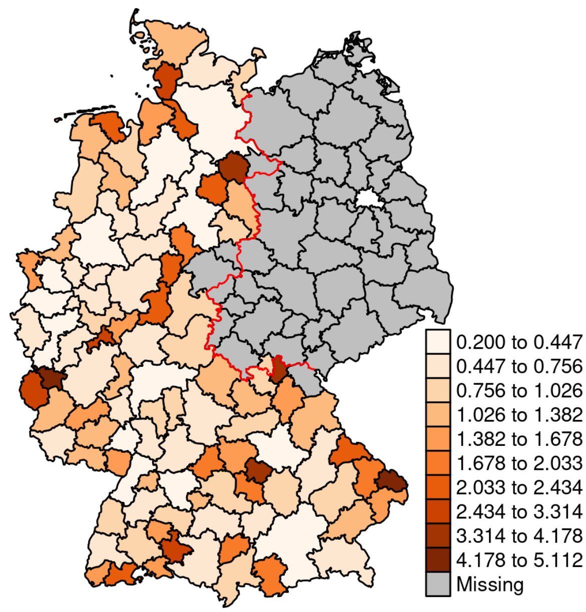
Table 4: Estimation of migration costs

	$n=R$		$n=L$	
	OLS	$PPML$	OLS	$PPML$
$\log(dist_{od})$	-1.001	-1.254	-1.063	-1.277
	(0.0147)	(0.0183)	(0.0208)	(0.0164)
ρ_1^n	0.4721	0.5915	0.5014	0.6023
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 (36). 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 $\log\left(\eta_{od}^n\right)$ 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 $\log(dist_{od})$, where $dist_{od}$ is the Euclidean distance in miles from o to d. Row 4 are standard errors two-way clustered at the o and d level. Row 5 is the implied migration elasticity from the estimates from row 3. Rows $6-7$ contain the goodness of fit and number of observations in each specification, respectively.

Location fundamentals $\{\mathcal{A}_o, \mathcal{B}_o^n\}_{\forall o\in\mathcal{S}}^{n=\{L,R\}}$. I follow Redding (2016) to recover unobserved fundamental location productivities $\{\mathcal{A}_o\}_{\forall o\in\mathcal{S}}$ and fundamental amenities for both workers and inventors $\{\mathcal{B}_o^n\}_{\forall o\in\mathcal{S}}^{n=\{L,R\}}$. First, I recover fundamental location productivities $\{\mathcal{A}_o\}_{\forall o\in\mathcal{S}}$. Given values for parameters $\{\sigma,\widetilde{\gamma}\}$, trade costs $\{\tau_{od}\}_{\forall o,d\in\mathcal{S},\mathcal{S}}$, and data on wages and population $\left\{w_o^L, w_o^R, L_o, R_o\right\}_{\forall o\in\mathcal{S}}$, there is a unique set of values for fundamental location productivities $\{\mathcal{A}_o\}_{\forall o\in\mathcal{S}}$ that is consistent with the data. Since the model is static, then I use data on wages and population from 2014 to denote that West Germany was at its steady-state equilibrium by then. To recover these fundamentals, I solve a fixed point algorithm on the system of excess demand functions implied by Equations (32). In Figure 1 I show the spatial distribution of these fundamentals. Notice that to rationalize the presence of production and innovation in less-dense locations, then these locations report higher fundamental levels of productivity.

Figure 1: Fundamental location productivity

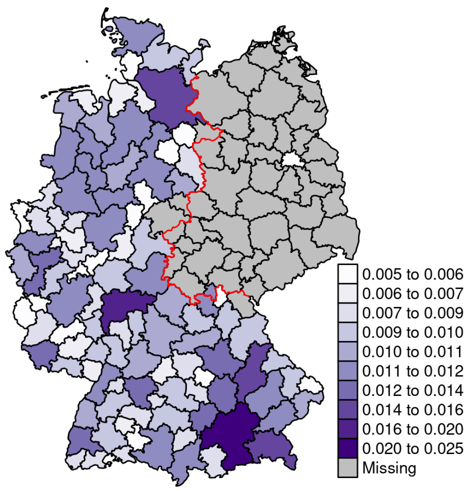


Notes: This figure shows the spatial distribution of fundamental location productivities $\{\mathcal{A}_o\}$ in West Germany. A darker (lighter) orange color denotes a higher (lower) productivity. All these values are normalized by their geometric mean.

Now, given values for parameters $\{\sigma, \kappa, \widetilde{\gamma}\}$, trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, migration costs $\{\mu^n_{od}\}^{n=\{L,R\}}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, fundamental location productivities $\{\mathcal{A}_o\}_{\forall o \in \mathcal{S}}$, and initial distribution of workers and inventors across locations $\left\{\overline{L}_o, \overline{R}_o\right\}_{\forall o \in \mathcal{S}}$, and data on wages and population $\left\{w^L_o, w^R_o, L_o, R_o\right\}_{\forall o \in \mathcal{S}}$, there is a unique set of values for fundamental location amenities $\{\mathcal{B}^n_o\}^{n=\{L,R\}}_{\forall o \in \mathcal{S}}$ that is consistent with the data. The data on wages and population is from 2014. The initial distribution $\left\{\overline{L}_o, \overline{R}_o\right\}_{\forall o \in \mathcal{S}}$ is from 1980 and they are scaled such that the total number of workers and inventors is the same for 2014. To recover these fundamentals, I solve a fixed point algorithm on the system of excess demand functions implied by Equations (29) and (27). In Figure 2 I show the spatial distribution of these fundamentals.

Figure 2: Fundamental location amenities

(a) Workers (b) Inventors



Notes: This figure shows the spatial distribution of fundamental location amenities $\{\mathcal{B}_o^L, \mathcal{B}_o^R\}$ in West Germany. A darker (lighter) purple color denotes a higher (lower) productivity. All these values are normalized by their corresponding arithmetic 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 $\xi_1 = \frac{1.56}{\sigma-1}$. I use this value instead of the one from Monte et al. (2018) since the former is based on internal trade data from Germany.

Rest of parameters $\{\alpha, \kappa, \sigma\}$. The remaining parameters are the dispersion of productivity of ideas α, the migration elasticity κ, and the elasticity of substitution across intermediate inputs σ. Regardless of the microfoundation for firms' R&D, a value of α is necessary to obtain values for the constant ψ from Equation (15). Following the process for the generation of ideas from Appendix D.1, 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).

For the migration elasticity κ, I follow Peters (2022) and set $\kappa = 2.12$. Using this value is appropriate for this setup since it is estimated using German data. Finally, for the elasticity of substitution, I follow Broda and Weinstein (2006) and set $\sigma = 2.5$. This value fits this setup since it is estimated for industrial sectors.

Table 5: Summary of calibration

	Description	Value	Identification/Moments
		Innovation	
$\widetilde{\gamma}$	Spatial 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
ζ	Decreasing returns to R&D	0.65	OLS estimate, Table 3, column 3
α	Idea productivity dispersion	1.5	Pareto parametric fit
		Migration	
$\{\rho_0^R, \rho_0^L\}$	Migration costs, intercepts	$\rho_0^R = 1.354$ $\rho_0^L = 1.361$	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	
$\mathcal{A}_o$	Location productivites		Recovered, Equation (32)
$\{\mathcal{B}_o^R, \mathcal{B}_o^L\}$	Location amenities		Recovered, Equations (27) and (29)
		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

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

4 Counterfactuals

In this section, I use the calibrated model to conduct two set of counterfactuals. First, I quantify the effect of reducing inventor migration costs by 25%. Second, I quantify the effect of the 2020 German R&D Tax Allowance Act which implemented a 25% subsidy for firms' R&D expenditure. In each counterfactual, I study the effect of these policies on aggregate productivity, and explore how these effects depend on spatial knowledge spillovers in R&D.

4.1 Reducing inventor migration costs

The model predicts that reducing inventor migration costs μ_{od}^R by 25% leads to an increase of aggregate productivity of 5.87%. This figure is comparable to Bryan and Morten (2019) who find that a 30% proportional reduction of both μ_{od}^{L} and μ_{od}^{R} lead to a 7% increase of aggregate output. Nevertheless, it is surprising that a similar reduction of μ_{od}^{R} can lead to comparable increases in aggregate productivity or output since inventors comprise a small share of the population. In that sense, fostering the mobility of the agents behind R&D could be a cost-effective way to promote economic activity.

Additionally, since the number of inventors in the economy is finite, this policy could generate heterogeneous effects across locations. In Figure 3 I analyze the productivity gains of each location, where these gains are measured as

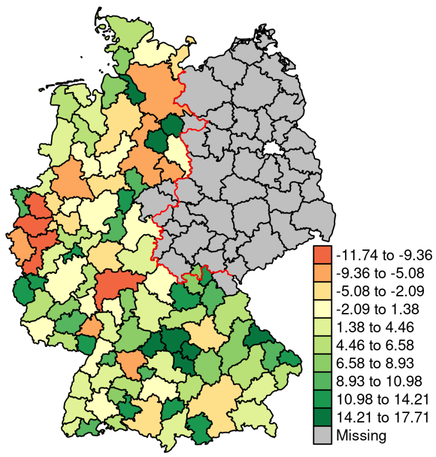
$$g_o^A = \left(\frac{A_o^{policy} - A_o^{baseline}}{A_o^{baseline}} \right) \times 100\%,$$

where A_o^{policy} is the productivity of location o under the reduction of migration costs for inventors, and $A_o^{baseline}$ is productivity of location o at the baseline scenario. The left panel of the figure shows a map with the location productivity gains of implementing the policy. The map shows that the policy indeed generates large heterogeneous effects across locations. At the upper tail of the distribution, there are locations that increase their productivity by around $14\% - 17\%$. In contrast, at the lower tail, some locations exhibit lower productivity by around $9\% - 11\%$.

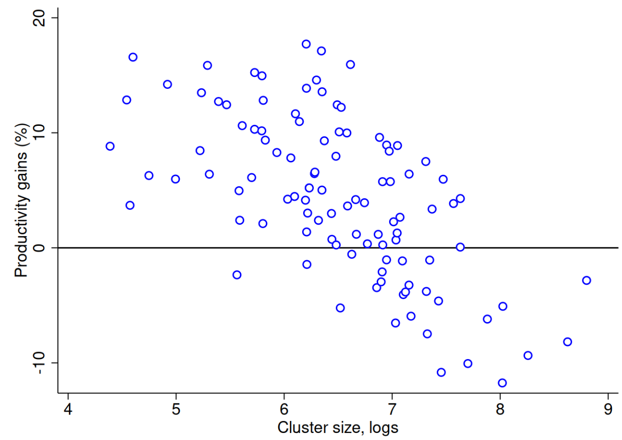
On the right panel of Figure 3, we observe that reducing μ_{od}^{R} exhibits an equalizing effect—larger clusters at the baseline exhibited decreases in productivity, while smaller clusters gained productivity—That is, after facilitating the spatial mobility of inventors, they prefer to move towards smaller clusters. For example, after the policy, a large share of inventors from Munich and Stuttgart moved to contiguous locations, so they exhibited the largest productivity gains. Inventors move towards smaller clusters after the policy because they can earn higher real wages.

Figure 3: Reduction of inventor migration costs by 25%

(a) Productivity gains, map



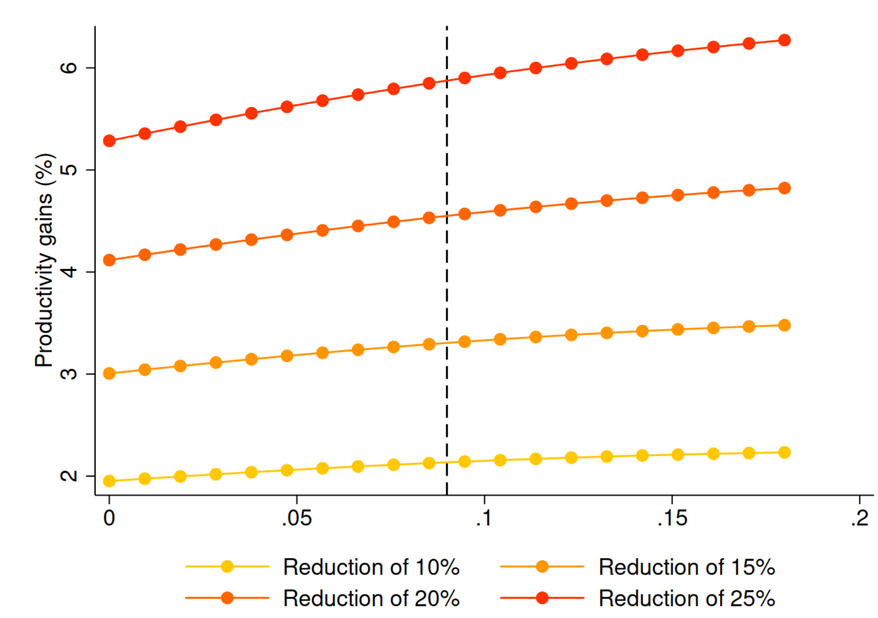
(b) Productivity gains VS cluster size in logs



Notes: This figure is comprised by two panels. I measure productivity gains as $g_o^A = \left(\frac{A_o^{policy} - A_o^{baseline}}{A_o^{baseline}}\right) \times 100\%$, where A_o^{policy} is aggregate productivity of location o under the 25% reduction of μ_{od}^R, and $A_o^{baseline}$ is aggregate productivity of location o at the baseline scenario. On the left panel, I color each location in West Germany according to their value of g_o^A. On the right panel, I compare g_o^A with the number of inventors in each location at the baseline $R_o^{baseline}$.

I now explore how spatial knowledge spillovers in R&D influence the effect of reducing μ_{od}^R on aggregate productivity. In Figure 4, intuitively, we observe that reducing μ_{od}^R unambiguously increases aggregate productivity. For example, consider the dashed vertical line for the calibrated value of $\widetilde{\gamma} = 0.09$. Then, larger reductions of μ_{od}^R exhibit larger productivity gains. More interestingly, the effect of reducing μ_{od}^R on aggregate productivity exhibits complementarity with the value of $\widetilde{\gamma}$. Considering a reduction of 10% (the yellow line), going from scenario of no spillovers ($\widetilde{\gamma} = 0$) to doubling the spillovers ($\widetilde{\gamma} = 0.09 \times 2$) generates additional 28*pp* productivity gains. In contrast, considering a reduction of 25% (the red line), the same exercise leads to additional 98*pp* productivity gains. This highlights the importance of implementing policies that foster both the mobility of inventors and spatial knowledge spillovers in R&D to promote economic activity.

Figure 4: Productivity gains VS $\widetilde{\gamma}$, by reduction of μ_{od}^{R}



Notes: This figure shows the relationship between productivity gains g^A of reducing μ_{od}^{R} VS spatial knowledge spillovers $\widetilde{\gamma}$. I measure productivity gains as $g^A = \left(\frac{A_o^{policy} - A_o^{baseline}}{A_o^{baseline}}\right) \times 100\%$, where A_o^{policy} is aggregate productivity of location o under the reduction of μ_{od}^{R}, and $A_o^{baseline}$ is aggregate productivity of location o at the baseline scenario. The horizontal axis exhibits different values of $\widetilde{\gamma}$, where the vertical dashed line indicates the calibrated value of $\widetilde{\gamma} = 0.09$. I plot four scenarios ranging from a reduction of μ_{od}^{R} of 10% (yellow) to a reduction of 25% (red).

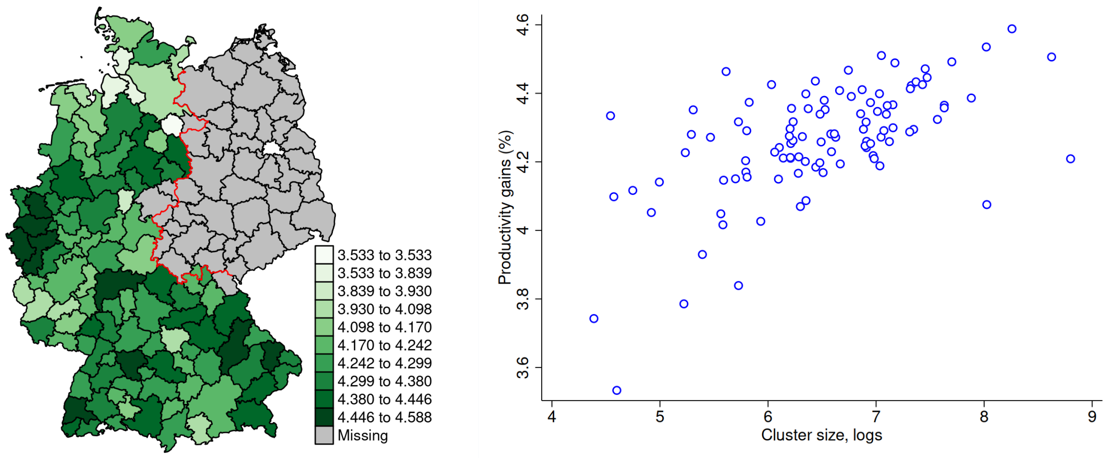
4.2 2020 German R&D Tax Allowance Act

In this section I 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). The model predicts that implementing an R&D subsidy $s_o = 25\%$ increases aggregate productivity by 4.27%. In Figure 5 we see that the implications for heterogeneity across locations is drastically different from the policy of reducing inventor migration costs. First, the left panel shows that the subsidy increases aggregate productivity everywhere, ranging from 3.5% to 4.5% gains. Second, the right panel shows that larger clusters are the ones that exhibited larger productivity gains since they increased even more in size, so the subsidy increased the spatial concentration of inventors.

Figure 5: Subsidy for firms' R&D expenditure by 25%

(a) Productivity gains, map

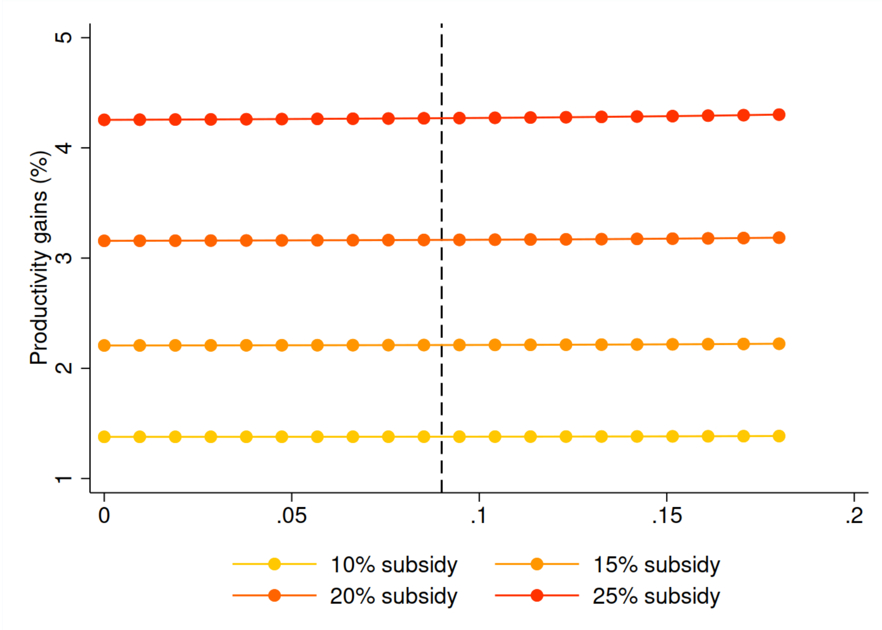
(b) Productivity gains VS cluster size in logs



Notes: This figure is comprised by two panels. I measure productivity gains as $g_o^A = \left(\frac{A_o^{policy} - A_o^{baseline}}{A_o^{baseline}}\right) \times 100\%$, where A_o^{policy} is aggregate productivity of location o under the R&D subsidy $s_o = 25\%$, and $A_o^{baseline}$ is aggregate productivity of location o at the baseline scenario. On the left panel, I color each location in West Germany according to their value of g_o^A. On the right panel, I compare g_o^A with the number of inventors in each location at the baseline $R_o^{baseline}$.

Now, I now explore how spatial knowledge spillovers in R&D influence the effect of implementing R&D subsidies on aggregate productivity. In Figure 6, intuitively, we observe that implemeting an R&D subsidy unambiguously increases aggregate productivity. In contrast with the policy of reducing inventor migration costs, the effect of implementing the subsidy on aggregate productivity exhibits weaker complementarity with the value of $\widetilde{\gamma}$. Considering a subsidy of 10% (the yellow line), going from scenario of no spillovers ($\widetilde{\gamma} = 0$) to doubling the spillovers ($\widetilde{\gamma} = 0.09 \times 2$) generates additional $0.7pp$ productivity gains. In contrast, considering a subsidy of 25% (the red line), the same exercise leads to additional $4.7pp$ productivity gains.

Figure 6: Productivity gains VS $\widetilde{\gamma}$, by value of s_o



Notes: This figure shows the relationship between productivity gains g^A of implementing an R&D subsidy $s_o = 25\%$ VS spatial knowledge spillovers $\widetilde{\gamma}$. I measure productivity gains as $g^A = \left(\frac{A_o^{policy} - A_o^{baseline}}{A_o^{baseline}}\right) \times 100\%$, where A_o^{policy} is aggregate productivity of location o under the R&D subsidy, and $A_o^{baseline}$ is aggregate productivity of location o at the baseline scenario. The horizontal axis exhibits different values of $\widetilde{\gamma}$, where the vertical dashed line indicates the calibrated value of $\widetilde{\gamma} = (0.409)(0.22) = 0.09$. I plot four scenarios ranging from $s_o = 10\%$ (yellow) to $s_o = 25\%$ (red).

5 Conclusions

In this paper I quantify the importance of spatial knowledge spillovers in R&D for aggregate productivity. 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 for aggregate productivity. I show that reducing migration costs for inventors and subsidies to firms' R&D activities can substantially increase aggregate productivity, and spatial 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 aggregate productivity. 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 spatial 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. The model is flexible enough to easily introduce these mechanisms, I leave these for future research.

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.

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.

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.

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.

Kortum, S. (1993). Equilibrium r&d and the patent–r&d ratio: Us evidence. *The American Economic Review*, 83(2):450–457.

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.

Krugman, P. R. (1979). Increasing returns, monopolistic competition, and international trade. *Journal of international Economics*, 9(4):469–479.

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.

Monte, F., Redding, S. J., and Rossi-Hansberg, E. (2018). Commuting, Migration, and Local Employment Elasticities. *American Economic Review*, 108(12):3855–90.

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.

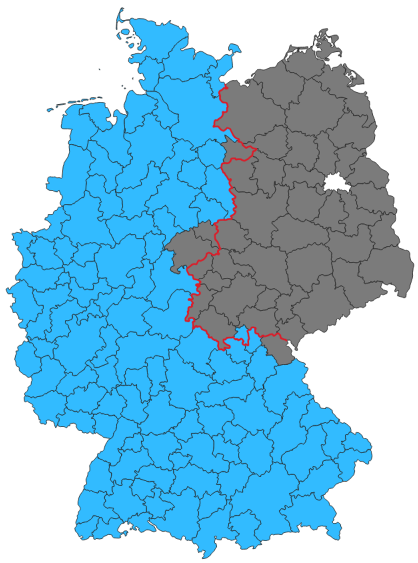
WEF (2018). The global competitiveness report 2018. World Economic Forum.

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

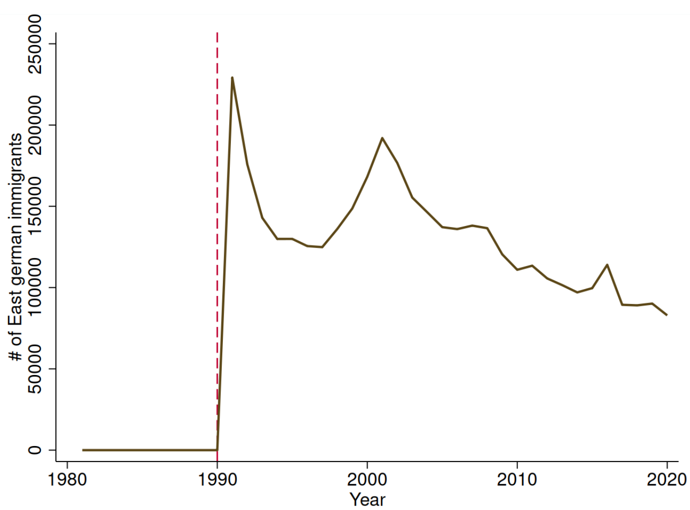
A Additional figures

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 B.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 B.2: OLS models, robustness

	(1)	(2)	(3)	(4)	(5)	(6)
	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 B.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 B.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 B.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 $\Delta \log\left(1+Z_{da,t}^{i\omega}\right)$, where $Z_{da,t}^{i\omega}$ 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 $\Delta IHS\left(Z_{da,t}^{i\omega}\right)$, 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 B.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
	(0.0107)	(0.0179)	(0.0254)	(0.0156)
ρ_1^n	0.4811	0.6514	0.7099	0.6509
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 (36) 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 $\log\left(\eta_{od}^n\right)$ is the log of the share of inventors or workers from o that moved to d during a given period, where I consider 1-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 and d level. Row 5 is the implied migration elasticity from the estimates from row 3. Rows $6-7$ contain the goodness of fit and number of observations in each specification, respectively.

Table B.7: Estimation of decreasing returns to R&D, 5-year periods

	(1)	(2)	(3)
$\log(R_o^{\omega})$	0.661	0.644	0.568
	(0.008)	(0.0081)	(0.0098)
ι_o		✓	✓
ι_ω			✓
N	95,699	95,699	68,381
R^2	0.683	0.744	0.854

Notes: In this table I report estimates for decreasing returns to R&D from Equation (35). To account for zeros on the dependent variable, I consider $\log(1 + n_o^{\omega})$. Time are 5-year periods. Each column is an specification with different combinations of fixed effects. The fixed effects included in each specification are determined by rows $4-5$. Row 2 contains the estimates for ζ, and row 3 contain standard errors, which are clustered at the ω level. Rows $6-7$ contain the number of observations and goodness of fit in each specification, respectively.

C Linked inventor biography data (INV-BIO)

The INV-BIO is comprised by around 150,000 inventors in Germany with high–frequency information on their employment spells and patenting activities between 1980 and 2014. All inventors recorded in the INV-BIO data filed at least one patent with the European Patent Office (EPO) between 1999 and 2011 and were disambiguated using a combination of record linkage and machine learning methods.[3] The INV-BIO dataset is comprised by three modules: (i) inventor module, (ii) establishment module, and (iii) patent module. I now describe the processing of each module.

C.1 Inventor module.

The module is reported at the employment spell level. The data is then collapsed at the inventor and period level. For a given inventor and period, consider the set of the inventor's spells. Then, for a given spell, the data contains information on the establishment an inventor works for, inventor's daily wage, 1-digit occupation code, whether job is part time, and the inventor's residence location. Since an inventor can have many jobs within a single period, I define the inventor's main job in the following manner. For a given inventor and period, the inventor's main job is the one with the longest tenure. Whenever a tie happens, the

[3] For more details, see `https://fdz.iab.de/en/FDZ_Individual_Data/INV-BIO-ADIAB/INV-BIO-ADIAB8014.aspx`

inventor's main job is the one with the highest daily wage. If ties still remain, an inventor's main job is chosen randomly. I exclude part-time jobs.

C.2 Establishment module.

The module is reported at the establishment and year level. The data is collapsed to the establishment and period level. The data contains a 1-digit 2008 time-consistent NACE code, the year the establishment appears for the first time in the administrative records, the year the establishment disappears from the administrative records, and the establishment's location. I construct a panel of establishments based on the years the establishments were first and last registered in administrative records.

C.3 Patent module.

The module is reported at the patent, inventor, and date-of-filing level. The data is collapsed to the inventor and period level. The data contains patent characteristics such as 2-10 forward year citations from the German Patent and Trade Mark Office (DPMA), the EPO, and the United States Patent and Trademark Office (US); the mean distance between the inventors that filed the patent, 1-digit technological area, and originality and generality indices. For each patent, the earliest filing date determines the date when the patent was generated. Then, for every inventor and period, I consider the set of patents filled by that inventor. I then collapse the data to the inventor and period level.

D Microfoundations

In this section I elaborate on the microfoundations of the model.

D.1 Generation of ideas

Consider an inventor i working for ω in location o. $Z_o^{i\omega,j}$ is the productivity of an idea j that inventor i generated. In this model, innovation is the process for which an inventor generates T_o ideas and selects the one with the highest productivity, such that

$$Z_o^{i\omega} = \max_{j=1,\ldots,T_o^{\omega}} Z_o^{i\omega,j}.$$

The conditional probability distribution of inventor i's best idea is

$$\begin{aligned}
G_o\left(z \mid T_o^\omega\right) &= \Pr\left\{Z_o^{i\omega,j} \leq z \mid T_o^\omega\right\}, \\
&= \Pr\left\{Z_o^{i\omega,1} \leq z, \ldots, Z_o^{i\omega,T_o^\omega} \leq z \mid T_o^\omega\right\}, \\
&= \Pr\left\{Z_o^{i\omega,1} \leq z\right\} \times \cdots \times \Pr\left\{Z_o^{i\omega,T_o^\omega} \leq z\right\}, \\
&= \underbrace{F\left(z\right) \times \cdots \times F\left(z\right)}_{T_o \ times}, \\
&= F\left(z\right)^{T_o},
\end{aligned}$$

where $F\left(z\right)$ is the cumulative probability that an idea drawn by inventor i is below productivity z. Since T_o is the discrete number of draws an inventor makes, I assume that T_o follows a Poisson distribution, such that $\Pr\left\{T_o = n\right\} = \frac{\lambda_o^n \exp(-\lambda_o)}{n!}$, where n is the number of drawn ideas, and λ_o is the expected number of drawn ideas. I also assume that ideas are drawn from a Pareto distribution $F\left(z\right) = 1 - z^{-\alpha}$, where $\alpha > 1$ is a shape parameter. Then, the unconditional distribution of the productivity of inventor i's best idea is

$$\begin{aligned}
G_o\left(z\right) &= \Pr\left\{Z_o^{i\omega} \leq z\right\}, \\
&= \sum_{n=0}^{\infty}\left[\frac{\lambda_o^n \exp\left(-\lambda_o\right)}{n!}\right]\left[F\left(z\right)^n\right], \\
&= \exp\left(-\lambda_o^\omega\right)\left[\sum_{n=0}^{\infty}\frac{\left(\lambda_o F\left(z\right)\right)^n}{n!}\right], \\
&= \exp\left(-\lambda_o\right)\exp\left(\lambda_o F\left(z\right)\right), \\
&= \exp\left(-\lambda_o\left(1 - F\left(z\right)\right)\right), \\
&= \exp\left(-\lambda_o\left(1 - \left(1 - z^{-\alpha}\right)\right)\right), \\
&= \exp\left(-\lambda_o z^{-\alpha}\right).
\end{aligned}$$

That is, $Z_o^{i\omega}$ is drawn from a Frechet distribution with shape parameter α and scale $\lambda_o^{\frac{1}{\alpha}}$.

D.2 Microfoundations: quality of intermediate inputs

D.2.1 Microfoundation 1: complementary tasks.

The blueprint of the intermediate input ω is comprised by a continuum of tasks $\mathcal{T} \equiv [0, 1]$, and all these tasks are necessary to produce the input at a certain quality. Then, the quality of the intermediate is

$$A_o^\omega = \exp\left(\int_{\mathcal{T}} \log\left(A_o^{\omega,\tau}\right) d\tau\right),$$

where $A_o^{\omega,\tau}$ is the quality of task $\tau \in \mathcal{T}$ within ω's blueprint. The firm hires a mass of inventors R_o^ω who generate $n_o^\omega \leq R_o^\omega$ ideas that determine the quality of each task within the ω's blueprint. Ideas are heterogeneous in productivity and each idea improves all tasks' quality. Then, the quality of each task is

$$A_o^{\omega,\tau} = z^\tau n_o^\omega,$$

where z^τ is the productivity of the ideas that inventors generated. Plugging this into the expression for A_o^ω yields

$$\begin{aligned}
A_o^\omega &= \exp\left(\int_{\mathcal{T}} \log\left(A_o^{\omega,\tau}\right) dt\right), \\
&= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau n_o^\omega\right) dt\right), \\
&= \exp\left(\int_{\mathcal{T}} \left[\log\left(z^\tau\right) + \log\left(n_o^\omega\right)\right] dt\right), \\
&= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau\right) d\tau + \int_{\mathcal{T}} \log\left(n_o^\omega\right) dt\right), \\
&= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau\right) d\tau\right) \exp\left(\int_{\mathcal{T}} \log\left(n_o^\omega\right) dt\right), \\
&= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau\right) d\tau\right) \exp\left(\log\left(n_o^\omega\right) \int_{\mathcal{T}} dt\right), \\
&= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau\right) d\tau\right) \exp\left(\log\left(n_o^\omega\right)\right), \\
&= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau\right) d\tau\right) n_o^\omega.
\end{aligned}$$

Since z^τ are drawn from a Frechet distribution as in Equation (13), then $\log\left(z^\tau\right)$ are drawn from a Gumbel distribution with location parameter $\log\left(\lambda_o^{\frac{1}{\alpha}}\right)$ and scale parameter $\frac{1}{\alpha}$. Then,

$$\begin{aligned}
A_o^\omega &= \exp\left(\int_{\mathcal{T}} \log\left(z^\tau\right) d\tau\right) n_o^\omega, \\
&= \exp\left(\int_0^\infty \log\left(z\right) dG_o\left(z\right)\right) n_o^\omega, \\
&= \exp\left(\log\left(\lambda_o^{\frac{1}{\alpha}}\right) + \frac{\overline{\gamma}}{\alpha}\right) n_o^\omega, \\
&= \exp\left(\log\left(\lambda_o^{\frac{1}{\alpha}}\right)\right) \exp\left(\frac{\overline{\gamma}}{\alpha}\right) n_o^\omega, \\
&= \psi \lambda_o^{\frac{1}{\alpha}} n_o^\omega.
\end{aligned}$$

where $\psi \equiv \exp\left(\frac{\overline{\gamma}}{\alpha}\right)$ is a constant, and $\overline{\gamma}$ is Euler's constant.

D.2.2 Microfoundation 2: linear innovation.

Consider a firm ω that hired a mass of inventors R_o^ω. The task of each inventor is to come up with an idea that will be incorporated into the input's blueprint. Inventors show up for work, they form an arbitrary line, the first inventor receives the blueprint, implements his idea into the blueprint and passes it over to the next inventor, and so on. At the end, $n_o^\omega \leq R_o^\omega$ ideas have been implemented into the blueprint since some inventors are not able to generate an idea due decreasing returns to R&D (e.g. duplication effects). Ideas are heterogeneous in productivity since they follow a Frechet distribution as in (13). Then, the quality of intermediate input ω under this microfoundation is

$$\begin{aligned} A_o^\omega &= \int_o^{n_o^\omega} z^i di, \\ &= n_o^\omega \int_o^\infty z dG_o(z), \\ &= n_o^\omega \left[\Gamma\left(1-\frac{1}{\alpha}\right)\lambda_o^{\frac{1}{\alpha}}\right], \\ &= \psi \lambda_o^{\frac{1}{\alpha}} n_o^\omega, \end{aligned}$$

where $\psi \equiv \Gamma\left(1-\frac{1}{\alpha}\right)$ is a constant, and $\Gamma(\cdot)$ is the Gamma function.

E Derivations

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

$$Q_d = \left(\sum_o \int_{\omega\in\Omega_{od}} A_o^{\omega\frac{1}{\sigma}} Q_{od}^{\omega\frac{\sigma-1}{\sigma}} d\omega\right)^{\frac{\sigma}{\sigma-1}}, \tag{37}$$

where Ω_{od} is the set of intermediate input firms in o selling to d, Q_d is the production of the final good, Q_{od}^ω is the quantity of intermediate input ω, A_o^ω is input quality, and $\sigma > 1$ is the

CES across intermediate inputs. The final good producer maximizes profits:

$$\max_{\{Q^{\omega}_{od}\}} \; P_d Q_d - \sum_o \int_{\omega\in\Omega_{od}} P^{\omega}_{od} Q^{\omega}_{od} s.t.$$
$$Q_d = \left(\sum_o \int_{\omega\in\Omega_{od}} A^{\omega\frac{1}{\sigma}}_{o} Q^{\omega\frac{\sigma-1}{\sigma}}_{od} \, d\omega \right)^{\frac{\sigma}{\sigma-1}}.$$

The first order condition of buying an intermediate input ω from o is

$$\begin{aligned}
[Q^{\omega}_{od}] : & P_d \left(\frac{\sigma}{\sigma-1}\right) (\ldots_d)^{\frac{\sigma}{\sigma-1}-1} A^{\omega\frac{1}{\sigma}}_{o} \left(\frac{\sigma-1}{\sigma}\right) Q^{\omega\frac{\sigma-1}{\sigma}-1}_{od} = P^{\omega}_{od}, \\
P^{\omega}_{od} &= P_d (\ldots_d)^{\frac{\sigma-\sigma+1}{\sigma-1}} A^{\omega\frac{1}{\sigma}}_{o} Q^{\omega\frac{\sigma-1-\sigma}{\sigma}}_{od}, \\
&= P_d (\ldots_d)^{\frac{1}{\sigma-1}} A^{\omega\frac{1}{\sigma}}_{o} Q^{\omega-\frac{1}{\sigma}}_{od},
\end{aligned}$$

where $(\ldots_d)$ is a composite of terms in d. Now, consider the first order condition of buying an intermediate input ω from o':

$$P^{\omega}_{od} = P_d (\ldots_d)^{\frac{1}{\sigma-1}} A^{\omega\frac{1}{\sigma}}_{o} Q^{\omega-\frac{1}{\sigma}}_{o'd}.$$

Divide both order conditions:

$$\begin{aligned}
\frac{P^{\omega}_{od}}{P^{\omega}_{o'd}} &= \frac{P_d (\ldots_d)^{\frac{1}{\sigma-1}} A^{\omega\frac{1}{\sigma}}_{o} Q^{\omega-\frac{1}{\sigma}}_{od}}{P_d (\ldots_d)^{\frac{1}{\sigma-1}} A^{\omega\frac{1}{\sigma}}_{o} Q^{\omega-\frac{1}{\sigma}}_{o'd}}, \\
&= \frac{A^{\omega\frac{1}{\sigma}}_{o} Q^{\omega-\frac{1}{\sigma}}_{od}}{A^{\omega\frac{1}{\sigma}}_{o'} Q^{\omega-\frac{1}{\sigma}}_{o'd}}, \\
&= \frac{A^{\omega\frac{1}{\sigma}}_{o} Q^{\omega\frac{1}{\sigma}}_{o'd}}{A^{\omega\frac{1}{\sigma}}_{o'} Q^{\omega\frac{1}{\sigma}}_{od}}, \\
\frac{P^{\omega\sigma-1}_{od}}{P^{\omega\sigma-1}_{o'd}} &= \frac{A^{\omega\frac{\sigma-1}{\sigma}}_{o} Q^{\omega\frac{\sigma-1}{\sigma}}_{o'd}}{A^{\omega\frac{\sigma-1}{\sigma}}_{o'} Q^{\omega\frac{\sigma-1}{\sigma}}_{od}}, \\
Q^{\omega\frac{\sigma-1}{\sigma}}_{o'd} &= \frac{A^{\omega\frac{\sigma-1}{\sigma}}_{o'} Q^{\omega\frac{\sigma-1}{\sigma}}_{od}}{A^{\omega\frac{\sigma-1}{\sigma}}_{o}} \frac{P^{\omega\sigma-1}_{od}}{P^{\omega\sigma-1}_{o'd}}.
\end{aligned}$$

Plug this expression in the production function of the final good producer:

$$
\begin{aligned}
Q_d &= \left(\sum_{o'} \int_{\omega \in \Omega_{o'd}} A_{o'}^{\omega\frac{1}{\sigma}} Q_{o'd}^{\omega\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \\
&= \left(\sum_{o'} \int_{\omega \in \Omega_{o'd}} A_{o'}^{\omega\frac{1}{\sigma}} \frac{A_{o'}^{\omega\frac{\sigma-1}{\sigma}} Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}{A_{o}^{\omega\frac{\sigma-1}{\sigma}}} \frac{P_{od}^{\omega\sigma-1}}{P_{o'd}^{\omega\sigma-1}} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \\
&= \left(\frac{Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}{A_{o}^{\omega\frac{\sigma-1}{\sigma}}} P_{od}^{\omega\sigma-1} \sum_{o'} \int_{\omega \in \Omega_{o'd}} A_{o'}^{\omega\frac{1}{\sigma}} A_{o'}^{\omega\frac{\sigma-1}{\sigma}} P_{o'd}^{\omega 1-\sigma} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \\
&= \left(\frac{Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}{A_{o}^{\omega\frac{\sigma-1}{\sigma}}} P_{od}^{\omega\sigma-1} \sum_{o'} \int_{\omega \in \Omega_{o'd}} A_{o'}^{\omega} P_{o'd}^{\omega 1-\sigma} d\omega \right)^{\frac{\sigma}{\sigma-1}}, \\
&= \left(\frac{Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}{A_{o}^{\omega\frac{\sigma-1}{\sigma}}} P_{od}^{\omega\sigma-1} \sum_{o'} P_{od,t}^{1-\sigma} \right)^{\frac{\sigma}{\sigma-1}}, \\
&= \left(\frac{Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}{A_{o}^{\omega\frac{\sigma-1}{\sigma}}} P_{od}^{\omega\sigma-1} P_{d,t}^{1-\sigma} \right)^{\frac{\sigma}{\sigma-1}},
\end{aligned}
$$

where $P_{d,t}^{1-\sigma} = \left(\sum_o P_{od,t}^{1-\sigma}\right)$, and $P_{od,t}^{1-\sigma} = \left(\int_{\omega \in \Omega_{od,t}} A_o^{\omega} P_{od}^{\omega 1-\sigma} d\omega\right)$ are CES price indices. Then, rearrange this expression to obtain the demand of intermediate inputs from o:

$$
\begin{aligned}
Q_d &= \left(\frac{Q_{od}^{\omega\frac{\sigma-1}{\sigma}}}{A_{o}^{\omega\frac{\sigma-1}{\sigma}}} P_{od}^{\omega\sigma-1} P_{d,t}^{1-\sigma} \right)^{\frac{\sigma}{\sigma-1}}, \\
&= \frac{Q_{od}^{\omega}}{A_o^{\omega}} P_{od}^{\omega\sigma} P_{d,t}^{-\sigma}, \\
Q_{od}^{\omega} &= A_o^{\omega} P_{od}^{\omega -\sigma} P_d^{\sigma} Q_d, \\
&= A_o^{\omega} P_{od}^{\omega -\sigma} P_d^{\sigma-1} \left(P_d Q_d\right), \\
&= A_o^{\omega} P_{od}^{\omega -\sigma} P_d^{\sigma-1} X_d,
\end{aligned}
$$

where $X_d = P_d Q_d$ is total expenditure of the final good in d.

Intermediate input firms. The intermediate input firm in o maximizes profits by selling its inputs to all locations subject to the demand from every location and its cost structure:

$$\begin{aligned}
\max_{\left\{P_{od}^{\omega},Q_{od}^{\omega},L_{od}^{\omega}\right\}} \pi_o^{\omega} &= \sum_d \pi_{od}^{\omega}, \\
&s.t. \\
\pi_{od} &= P_{od}^{\omega} Q_{od}^{\omega} - \tau_{od} w_o^L L_{od}^{\omega}, \\
L_{od}^{\omega} &= Q_{od}^{\omega}, \\
Q_{od}^{\omega} &= A_o^{\omega} P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d.
\end{aligned}$$

Introduce the constraints into the profit function:

$$\begin{aligned}
\pi_o^{\omega} &= \sum_d \pi_{od}^{\omega}, \\
&= \sum_d \left(P_{od}^{\omega} Q_{od}^{\omega} - \tau_{od} w_o^L L_{od}^{\omega}\right), \\
&= \sum_d \left(P_{od}^{\omega} Q_{od}^{\omega} - \tau_{od} w_o^L Q_{od}^{\omega}\right), \\
&= \sum_d \left(P_{od}^{\omega} A_o^{\omega} P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d\right) \\
&- \sum_d \left(\tau_{od} w_o^L A_o^{\omega} P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d\right), \\
&= \sum_d \left(A_o^{\omega} P_{od}^{\omega^{1-\sigma}} P_d^{\sigma-1} X_d\right) \\
&- \sum_d \left(\tau_{od} w_o^L A_o^{\omega} P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d\right).
\end{aligned}$$

The first order condition is

$$\begin{aligned}
\left[P_{od}^{\omega}\right] : (1-\sigma) A_o^{\omega} P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d - (-\sigma) \tau_{od} w_o^L A_o^{\omega} P_{od}^{\omega^{-\sigma-1}} P_d^{\sigma-1} X_d &= 0, \\
0 = (1-\sigma) P_{od}^{\omega^{-\sigma}} + \sigma \tau_{od} w_o^L P_{od}^{\omega^{-\sigma-1}}&, \\
(\sigma - 1) = \sigma \tau_{od} w_o^L P_{od}^{\omega^{-1}}&, \\
P_{od}^{\omega} = \left(\frac{\sigma}{\sigma-1}\right) \tau_{od} w_o^L&, \\
= \overline{m} \tau_{od} w_o^L&,
\end{aligned}$$

where $\overline{m} \equiv \frac{\sigma}{\sigma-1}$ is the CES constant markup over marginal costs.

Total profits. Introducing the markup pricing Equation (9) in the profit function (8) yields

$$
\begin{aligned}
\pi_o^\omega &= \sum_d \pi_{od}^\omega, \\
&= \sum_d P_{od}^\omega Q_{od}^\omega - \tau_{od} w_o^L L_{od}^\omega, \\
&= \sum_d P_{od}^\omega Q_{od}^\omega - \tau_{od} w_o^L Q_{od}^\omega, \\
&= \sum_d \left(P_{od}^\omega - \tau_{od} w_o^L\right) Q_{od}^\omega, \\
&= \sum_d \left(\overline{m}\tau_{od} w_o^L - \tau_{od} w_o^L\right) Q_{od}^\omega, \\
&= \sum_d \left(\overline{m} - 1\right) \tau_{od} w_o^L Q_{od}^\omega, \\
&= \sum_d \left(\overline{m} - 1\right) \tau_{od} w_o^L A_o^\omega P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d, \\
&= \sum_d \left(\overline{m} - 1\right) \tau_{od} w_o^L A_o^\omega \left(\overline{m}\tau_{od} w_o^L\right)^{-\sigma} P_d^{\sigma-1} X_d, \\
&= \sum_d \left(\overline{m} - 1\right) \overline{m}^{-1} \left(\overline{m}\tau_{od} w_o^L\right) A_o^\omega \left(\overline{m}\tau_{od} w_o^L\right)^{-\sigma} P_d^{\sigma-1} X_d, \\
&= A_o^\omega \sum_d \left(\overline{m} - 1\right) \overline{m}^{-1} \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma} P_d^{\sigma-1} X_d, \\
&= \left(\frac{\overline{m} - 1}{\overline{m}}\right) A_o^\omega \sum_d \left(P_{od}^\omega\right)^{1-\sigma} P_d^{\sigma-1} X_d, \\
&= \left(\frac{\frac{\sigma}{\sigma-1} - 1}{\frac{\sigma}{\sigma-1}}\right) A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d, \\
&= \left(\frac{\frac{1}{\sigma-1}}{\frac{\sigma}{\sigma-1}}\right) A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d, \\
&= \frac{1}{\sigma} A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d.
\end{aligned}
$$

Research and Development (R&D). Firm ω maximizes total profits after R&D expenditure subject to the quality of its intermediate:

$$\max_{\{R_o^\omega\}} \overline{\pi}_o^\omega = \pi_o^\omega - w_o^R R_o^\omega$$
$$s.t.$$
$$\pi_o^\omega = \frac{1}{\sigma} A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d,$$
$$A_o^\omega = \psi \mathcal{A}_o R_o^{\widetilde{\gamma}} R_o^{\omega\,\zeta}.$$

Rewrite profits:

$$\begin{aligned}
\overline{\pi}_o^\omega &= \pi_o^\omega - w_o^R R_o^\omega, \\
&= \frac{1}{\sigma} A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d - w_o^R R_o^\omega, \\
&= \frac{1}{\sigma} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right) n_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d - w_o^R R_o^\omega, \\
&= \frac{1}{\sigma} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right) R_o^{\omega\,\zeta} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d - w_o^R R_o^\omega.
\end{aligned}$$

The first order condition is

$$\begin{aligned}
[R_o^\omega] : w_o^R &= \frac{\zeta}{\sigma} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right) R_o^{\omega\,\zeta-1} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d, \\
R_o^{\omega\,1-\zeta} &= \frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d, \\
R_o^\omega &= \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}.
\end{aligned}$$

The number of implemented ideas is

$$\begin{aligned}
n_o^\omega &= R_o^{\omega\,\zeta}, \\
&= \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{\zeta}{1-\zeta}}.
\end{aligned}$$

The quality of the intermediate is

$$
\begin{aligned}
A_o^{\omega} &= \psi \mathcal{A}_o R_o^{\widetilde{\gamma}} n_o^{\omega}, \\
&= \psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \sum_d \left(\frac{P_{od}^{\omega}}{P_d} \right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right)^{\frac{1-\zeta}{\zeta}} \sum_d \left(\frac{P_{od}^{\omega}}{P_d} \right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\frac{\zeta}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right)^{1+\frac{1-\zeta}{\zeta}}}{w_o^R} \sum_d \left(\frac{P_{od}^{\omega}}{P_d} \right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\frac{\zeta}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right)^{\frac{\zeta+1-\zeta}{\zeta}}}{w_o^R} \sum_d \left(\frac{P_{od}^{\omega}}{P_d} \right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\frac{\zeta}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right)^{\frac{1}{\zeta}}}{w_o^R} \sum_d \left(\frac{P_{od}^{\omega}}{P_d} \right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\frac{\zeta}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right)^{\frac{1}{\zeta}}}{w_o^R} \sum_d \left(\frac{\overline{m} \tau_{od} w_o^L}{P_d} \right)^{1-\sigma} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \right)^{\frac{1}{\zeta}}}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{\zeta}{1-\zeta}}.
\end{aligned}
$$

Then, total profits after R&D is

$$
\begin{aligned}
\overline{\pi}_o^\omega &= \frac{1}{\sigma} A_o^\omega \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d - w_o^R R_o^\omega, \\
&= \frac{1}{\sigma} \left(\frac{\zeta}{\sigma} \frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{\zeta}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{\zeta}{1-\zeta}} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d \\
&- w_o^R \left(\frac{\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}, \\
&= \frac{1}{\sigma} \left(\frac{\zeta}{\sigma} w_o^{R^{-1}}\right)^{\frac{\zeta}{1-\zeta}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{1+\frac{\zeta}{1-\zeta}} \\
&- w_o^R \left(\frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R}\right)^{\frac{1}{1-\zeta}} \left(\frac{\zeta}{\sigma} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}, \\
&= \frac{1}{\sigma} \left(\frac{\zeta}{\sigma}\right)^{\frac{\zeta}{1-\zeta}} \left(w_o^R\right)^{\frac{\zeta}{\zeta-1}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}} \\
&- w_o^R \left(w_o^R\right)^{-\frac{1}{1-\zeta}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\frac{\zeta}{\sigma} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}, \\
&= \frac{\zeta^{\frac{\zeta}{1-\zeta}}}{\sigma^{1+\frac{\zeta}{1-\zeta}}} \left(w_o^R\right)^{\frac{\zeta}{\zeta-1}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}} \\
&- \left(\frac{\zeta}{\sigma}\right)^{\frac{1}{1-\zeta}} \left(w_o^R\right)^{1-\frac{1}{1-\zeta}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}, \\
&= \frac{\zeta^{\frac{\zeta}{1-\zeta}}}{\sigma^{\frac{1}{1-\zeta}}} \left(w_o^R\right)^{\frac{\zeta}{\zeta-1}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}} \\
&- \frac{\zeta^{\frac{1}{1-\zeta}}}{\sigma^{\frac{1}{1-\zeta}}} \left(w_o^R\right)^{\frac{\zeta}{\zeta-1}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}, \\
&= \left(\zeta^{\frac{\zeta}{1-\zeta}} - \zeta^{\frac{1}{1-\zeta}}\right) \left(w_o^R\right)^{\frac{\zeta}{\zeta-1}} \left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{1-\zeta}} \left(\frac{1}{\sigma} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}, \\
&= \kappa_\zeta^{\frac{1}{1-\zeta}} \left(w_o^{R^{-\zeta}} \psi \mathcal{A}_o R_o^{\widetilde{\gamma}} \frac{1}{\sigma} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}, \\
&= \left(\frac{\kappa_\zeta}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R^\zeta}} \sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}, \\
&= \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R^\zeta} w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d\right)^{\frac{1}{1-\zeta}}.
\end{aligned}
$$

where $\kappa_\zeta \equiv (1-\zeta)^{1-\zeta}\zeta^\zeta$ is a normalization constant. The normalization constant is

$$\begin{aligned}
\zeta^{\frac{\zeta}{1-\zeta}} - \zeta^{\frac{1}{1-\zeta}} &= \zeta^{\frac{\zeta}{1-\zeta}} - \zeta^{\frac{1}{1-\zeta}}, \\
&= \zeta^{\frac{1}{1-\zeta}} \left[\zeta^{\frac{\zeta}{1-\zeta}-\frac{1}{1-\zeta}} - 1\right], \\
&= \zeta^{\frac{1}{1-\zeta}} \left[\zeta^{\frac{\zeta-1}{1-\zeta}} - 1\right], \\
&= \zeta^{\frac{1}{1-\zeta}} \left[\zeta^{-\frac{1-\zeta}{1-\zeta}} - 1\right], \\
&= \zeta^{\frac{1}{1-\zeta}} \left[\frac{1}{\zeta} - 1\right], \\
&= \zeta^{\frac{1}{1-\zeta}} \left[\frac{1-\zeta}{\zeta}\right], \\
&= \zeta^{\frac{1}{1-\zeta}-1} (1-\zeta), \\
&= \zeta^{\frac{1-1+\zeta}{1-\zeta}} (1-\zeta), \\
&= (1-\zeta)\,\zeta^{\frac{\zeta}{1-\zeta}}, \\
&= \left[(1-\zeta)\,\zeta^{\frac{\zeta}{1-\zeta}}\right]^{\frac{1-\zeta}{1-\zeta}}, \\
&= \left[(1-\zeta)^{1-\zeta}\,\zeta^{\zeta}\right]^{\frac{1}{1-\zeta}}, \\
&= \kappa_\zeta^{\frac{1}{1-\zeta}}.
\end{aligned}$$

R&D with subsidies. Firm's demand for inventors is

$$R_o^\omega = \left(\frac{\zeta}{\sigma}\frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{(1-s_o)\,w_o^R}\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{1}{1-\zeta}}.$$

The number of ideas is

$$n_o^\omega = \left(\frac{\zeta}{\sigma}\frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{(1-s_o)\,w_o^R}\sum_d \left(\frac{P_{od}^\omega}{P_d}\right)^{1-\sigma} X_d\right)^{\frac{\zeta}{1-\zeta}}.$$

Intermediate's quality is

$$A_o^\omega = \left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma}\frac{\left(\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{\zeta}}}{(1-s_o)\,w_o^R w_o^{L^{\sigma-1}}}\sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d\right)^{\frac{\zeta}{1-\zeta}}.$$

Firms' profits are

$$\overline{\pi}_o^{\omega} = \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{\left(1-s_o\right)^{\zeta} w_o^{R\zeta} w_o^{L\sigma-1}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}.$$

Preferences. In each location d, agents are of two types: inventors ($n = R$), or workers ($n = L$). Each agent has preferences over the local final good and location amenities:

$$\begin{aligned} \max_{\{Q_d^n\}} \quad & U_d^n = \mathcal{B}_d^n Q_d^n \quad s.t. \\ & w_d^n \left(1+\overline{\pi}\right) = P_d Q_d^n, \end{aligned}$$

where $\mathcal{B}_d^n$ are type-specific location amenities, Q_d^n is the quantity demanded by agent of type n, w_d^n is the agent's wage, and $\overline{\pi}$ are redistributed profits. Since the agent's preferences are linear, utility is maximized at

$$\begin{aligned} U_d^n &= \mathcal{B}_d^n Q_d^n, \\ &= \mathcal{B}_d^n \left(\frac{w_d^n \left(1+\overline{\pi}\right)}{P_d} \right), \\ &= \frac{\mathcal{B}_d^n w_d^n \left(1+\overline{\pi}\right)}{P_d}. \end{aligned}$$

Location choice. An agent i of type $n = \{L, R\}$ living in o moves to d by maximizing its *ex ante* indirect utility:

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

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

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

such that $\sum_d \eta_{od}^n = 1, \forall o \in \mathcal{S}$.

Aggregate productivity. I define aggregate productivity as the average quality of intermediates in a location. Since firms are symmetric, then from equations (14)-(15), location's productivity is

$$
\begin{aligned}
A_o &= \int_{\omega\in\Omega_o} A_o^\omega d\omega, \\
&= \int_{\omega\in\Omega_o} \left(\frac{\zeta\overline{m}^{1-\sigma}}{\sigma} \frac{\left(\psi\mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{\zeta}}}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{\zeta}{1-\zeta}} d\omega, \\
&= \left(\frac{\zeta\overline{m}^{1-\sigma}}{\sigma} \frac{\left(\psi\mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{\zeta}}}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{\zeta}{1-\zeta}} \int_{\omega\in\Omega_o} d\omega, \\
&= \left(\frac{\zeta\overline{m}^{1-\sigma}}{\sigma} \frac{\left(\psi\mathcal{A}_o R_o^{\widetilde{\gamma}}\right)^{\frac{1}{\zeta}}}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{\zeta}{1-\zeta}}, \\
&= \left(\left(\frac{\zeta\overline{m}^{1-\sigma}}{\sigma} \right)^{\zeta} \left(\psi\mathcal{A}_o R_o^{\widetilde{\gamma}}\right) \left(w_o^{L^{\sigma-1}} w_o^R \right)^{-\zeta} \left(\sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\zeta} \right)^{\frac{1}{1-\zeta}}.
\end{aligned}
$$

With R&D subsidies, this is

$$
A_o = \left(\left(\frac{\zeta\overline{m}^{1-\sigma}}{\sigma} \right)^{\zeta} \left(\psi\mathcal{A}_o R_o^{\widetilde{\gamma}}\right) \left((1-s_o)\, w_o^{L^{\sigma-1}} w_o^R \right)^{-\zeta} \left(\sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\zeta} \right)^{\frac{1}{1-\zeta}}.
$$

Then, aggregate productivity is

$$
A = \frac{1}{S} \sum_o A_o,
$$

where $S \equiv |\mathcal{S}|$ is the number of locations in the economy.

Price indices. Given Equation (9), the price index of firms in o selling to d is

$$
\begin{aligned}
P_{od}^{1-\sigma} &= \int_{\omega\in\Omega_{od}} A_o^\omega P_{od}^{\omega^{1-\sigma}} d\omega, \\
&= \int_{\omega\in\Omega_{od}} A_o^\omega \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma} d\omega, \\
&= \int_{\omega\in\Omega_{od}} A_o^\omega d\omega \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma}, \\
&= \int_{\omega\in\Omega_{o}} A_o^\omega d\omega \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma}, \\
&= A_o \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma}.
\end{aligned}
$$

Then, the price index in d is

$$
\begin{aligned}
P_d^{1-\sigma} &= \sum_o P_{od}^{1-\sigma}, \\
&= \sum_o A_o \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma}.
\end{aligned}
$$

Trade shares. From (5), trade flows from o to d are

$$
\begin{aligned}
Q_{od}^{\omega} &= A_o^{\omega} P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d, \\
P_{od}^{\omega} Q_{od}^{\omega} &= A_o^{\omega} P_{od}^{\omega^{1-\sigma}} P_d^{\sigma-1} X_d, \\
X_{od}^{\omega} &= A_o^{\omega} P_{od}^{\omega^{1-\sigma}} P_d^{\sigma-1} X_d, \\
\int_{\omega\in\Omega_{od}} X_{od}^{\omega} d\omega &= \int_{\omega\in\Omega_{od}} A_o^{\omega} P_{od}^{\omega^{1-\sigma}} P_d^{\sigma-1} X_d d\omega, \\
X_{od} &= \left(\int_{\omega\in\Omega_{od}} A_o^{\omega} P_{od}^{\omega^{1-\sigma}} d\omega\right) P_d^{\sigma-1} X_d, \\
&= P_{od}^{1-\sigma} P_d^{\sigma-1} X_d,
\end{aligned}
$$

where $P_{od}^{1-\sigma} = \int_{\omega\in\Omega_{od}} A_o^{\omega} P_{od}^{\omega^{1-\sigma}} d\omega$. Then, the share of intermediate inputs from o in location d's expenditure χ_{od} is

$$
\begin{aligned}
\chi_{od} &\equiv \frac{X_{od}}{X_d}, \\
&= \frac{P_{od}^{1-\sigma} P_d^{\sigma-1} X_d}{X_d}, \\
&= P_{od}^{1-\sigma} P_d^{\sigma-1}, \\
&= \frac{P_{od}^{1-\sigma}}{P_d^{1-\sigma}}.
\end{aligned}
$$

Considering Equations (22), then trade shares are

$$
\begin{aligned}
\chi_{od} &= \frac{P_{od}^{1-\sigma}}{P_d^{1-\sigma}}, \\
&= \frac{A_o \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma}}{\sum_o A_o \left(\overline{m}\tau_{od} w_o^L\right)^{1-\sigma}}, \\
&= \frac{A_o \left(\tau_{od} w_o^L\right)^{1-\sigma}}{\sum_o A_o \left(\tau_{od} w_o^L\right)^{1-\sigma}}.
\end{aligned}
$$

Profits per-capita. Location's profits are

$$
\begin{aligned}
\overline{\pi}_o &= \int_{\omega\in\Omega_o} \overline{\pi}_o^{\omega} d\omega, \\
&= \int_{\omega\in\Omega_o} \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R^\zeta} w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}} d\omega, \\
&= \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R^\zeta} w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}} \int_{\omega\in\Omega_o} d\omega, \\
&= \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma} \psi}{\sigma} \frac{\mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R^\zeta} w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}.
\end{aligned}
$$

With R&D subsidies, these are

$$
\overline{\pi}_o = \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma} \psi}{\sigma} \frac{\mathcal{A}_o R_o^{\widetilde{\gamma}}}{(1-s_o)^\zeta \, w_o^{R^\zeta} w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}.
$$

Then, profits per-capita are

$$
\overline{\pi} = \frac{1}{N} \sum_o \pi_o,
$$

where $N = L + R$ is total population.

Workers market. From (5), (7), and (9), the demand for workers is

$$\begin{aligned}
L_o &= \sum_d \int_{\omega\in\Omega_{od}} L_{od}^{\omega} d\omega, \\
&= \sum_d \int_{\omega\in\Omega_{od}} Q_{od}^{\omega} d\omega, \\
&= \sum_d \int_{\omega\in\Omega_{od}} \left(A_o^{\omega} P_{od}^{\omega^{-\sigma}} P_d^{\sigma-1} X_d \right) d\omega, \\
&= \sum_d \int_{\omega\in\Omega_{od}} \left(A_o^{\omega} \left(\overline{m}\tau_{od} w_o^L\right)^{-\sigma} P_d^{\sigma-1} X_d \right) d\omega, \\
&= \sum_d \left(\overline{m}\tau_{od} w_o^L\right)^{-\sigma} P_d^{\sigma-1} X_d \int_{\omega\in\Omega_{od}} A_o^{\omega} d\omega, \\
&= \sum_d \left(\overline{m}\tau_{od} w_o^L\right)^{-\sigma} P_d^{\sigma-1} X_d \int_{\omega\in\Omega_{o}} A_o^{\omega} d\omega, \\
&= \sum_d \left(\overline{m}\tau_{od} w_o^L\right)^{-\sigma} P_d^{\sigma-1} X_d A_o, \\
&= A_o \left(\overline{m} w_o^L\right)^{-\sigma} \sum_d \tau_{od}^{-\sigma} P_d^{\sigma-1} X_d, \\
w_o^{L^{\sigma}} &= \frac{A_o}{L_o} \overline{m}^{-\sigma} \sum_d \tau_{od}^{-\sigma} P_d^{\sigma-1} X_d, \\
w_o^L &= \overline{m}^{-1} \left(\frac{A_o}{L_o} \sum_d \tau_{od}^{-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{\sigma}}.
\end{aligned}$$

Inventors market. From firm's demand for inventors (17), location's demand for inventors is

$$
\begin{aligned}
R_o^\omega &= \left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}, \\
\int_{\omega \in \Omega_o} R_o^\omega d\omega &= \int_{\omega \in \Omega_o} \left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}} d\omega, \\
R_o &= \left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}} \int_{\omega \in \Omega_o} d\omega, \\
&= R_o^{\frac{\widetilde{\gamma}}{1-\zeta}} \left(\frac{\zeta \overline{m}^{1-\sigma}}{\sigma} \frac{\psi \mathcal{A}_o}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}, \\
R_o^{1-\frac{\widetilde{\gamma}}{1-\zeta}} &= \left(\frac{\zeta \psi \overline{m}^{1-\sigma}}{\sigma} \frac{\mathcal{A}_o}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}, \\
R_o^{\frac{(1-\zeta)-\widetilde{\gamma}}{1-\zeta}} &= \left(\frac{\zeta \psi \overline{m}^{1-\sigma}}{\sigma} \frac{\mathcal{A}_o}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d \right)^{\frac{1}{1-\zeta}}, \\
R_o^{(1-\zeta)-\widetilde{\gamma}} &= \frac{\zeta \psi \overline{m}^{1-\sigma}}{\sigma} \frac{\mathcal{A}_o}{w_o^R w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d, \\
w_o^R &= \frac{\zeta \psi \overline{m}^{1-\sigma}}{\sigma} \frac{\mathcal{A}_o R_o^{\widetilde{\gamma}-(1-\zeta)}}{w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d.
\end{aligned}
$$

With R&D subsidies, this is

$$
w_o^R = \frac{\zeta \psi \overline{m}^{1-\sigma}}{\sigma} \frac{\mathcal{A}_o R_o^{\widetilde{\gamma}-(1-\zeta)}}{(1-s_o)\, w_o^{L^{\sigma-1}}} \sum_d \tau_{od}^{1-\sigma} P_d^{\sigma-1} X_d.
$$

F Taking the Model to the Data

Spatial knowledge spillovers in R&D $\{\widetilde{\gamma}\}$. From Equation (13), inventor productivity can be written as

$$\begin{aligned}
\mathbb{E}\left\{Z_o^{i\omega}\right\} &= \Gamma\left(1-\frac{1}{\alpha}\right)\lambda_o^{\frac{1}{\alpha}}, \\
&= \Gamma\left(1-\frac{1}{\alpha}\right)\left(\mathcal{A}_o R_o^{\gamma}\right)^{\frac{1}{\alpha}}, \\
&= \Gamma\left(1-\frac{1}{\alpha}\right)\mathcal{A}_o^{\frac{1}{\alpha}} R_o^{\frac{\gamma}{\alpha}}, \\
&= \Gamma\left(1-\frac{1}{\alpha}\right)\mathcal{A}_o^{\frac{1}{\alpha}} R_o^{\widetilde{\gamma}}, \\
Z_o^{i\omega} &= \Gamma\left(1-\frac{1}{\alpha}\right)\mathcal{A}_o^{\frac{1}{\alpha}} R_o^{\widetilde{\gamma}} \exp\left(\epsilon_o^{i\omega}\right), \\
\log\left(Z_o^{i\omega}\right) &= \log\left(\Gamma\left(1-\frac{1}{\alpha}\right)\mathcal{A}_o^{\frac{1}{\alpha}} R_o^{\widetilde{\gamma}} \exp\left(\epsilon_o^{i\omega}\right)\right), \\
&= \log\left(\Gamma\left(1-\frac{1}{\alpha}\right)\right) + \log\left(\mathcal{A}_o^{\frac{1}{\alpha}}\right) + \log\left(R_o^{\widetilde{\gamma}}\right) + \log\left(\exp\left(\epsilon_o^{i\omega}\right)\right), \\
&= \log\left(\Gamma\left(1-\frac{1}{\alpha}\right)\right) + \frac{1}{\alpha}\log\left(\mathcal{A}_o\right) + \widetilde{\gamma}\log\left(R_o\right) + \epsilon_o^{i\omega}, \\
&= \iota + \iota_o + \widetilde{\gamma}\log\left(R_o\right) + \epsilon_o^{i\omega},
\end{aligned}$$

where $\iota \equiv \log\left(\Gamma\left(1-\frac{1}{\alpha}\right)\right)$ is an intercept, and $\iota_o \equiv \frac{1}{\alpha}\log\left(\mathcal{A}_o\right)$ is location-specific fixed effect. In first-differences, this expression becomes

$$\begin{aligned}
\log\left(Z_o^{i\omega}\right) &= \iota + \iota_o + \widetilde{\gamma}\log\left(R_o\right) + \epsilon_o^{i\omega}, \\
\Delta\log\left(Z_o^{i\omega}\right) &= \Delta\left(\iota + \iota_o + \widetilde{\gamma}\log\left(R_o\right) + \epsilon_o^{i\omega}\right), \\
&= \Delta\iota + \Delta\iota_o + \widetilde{\gamma}\Delta\log\left(R_o\right) + \Delta\epsilon_o^{i\omega}, \\
&= \widetilde{\gamma}\Delta\log\left(R_o\right) + \Delta\epsilon_o^{i\omega}.
\end{aligned}$$

This expression is the model counterpart of Equation 2 that was used to estimate spatial knowledge spillovers in R&D.

Migration costs $\{\mu_{od}^n\}$. Migration costs are fully parametrized by data on geographic distance between every location pair $o, d \in \mathcal{S}, \mathcal{S}$, intercepts $\left\{\rho_o^L, \rho_o^R\right\}$, and elasticities of migration costs to distance $\left\{\rho_1^L, \rho_1^R\right\}$. Distances are constructed as described in Section 1.2. The intercepts are calibrated by targeting the overall migration rate for workers and inventors.

To estimate the migration cost elasticities, consider the share of inventors in Equation (20), such that

$$
\begin{aligned}
\eta_{od}^{n} &= \frac{\left(\frac{U_d^n}{\mu_{od}^n}\right)^{\kappa}}{\sum_{\delta}\left(\frac{U_\delta^n}{\mu_{o\delta}^R}\right)^{\kappa}}, \\
\log\left(\eta_{od}^{n}\right) &= \log\left(\frac{\left(\frac{U_d^n}{\mu_{od}^n}\right)^{\kappa}}{\sum_{\delta}\left(\frac{U_{\delta,t}^n}{\mu_{o\delta}^n}\right)^{\kappa}}\right), \\
&= \log\left(\left(\frac{U_d^n}{\mu_{od}^n}\right)^{\kappa}\right) - \log\left(\sum_{\delta}\left(\frac{U_\delta^n}{\mu_{o\delta}^n}\right)^{\kappa}\right), \\
&= \kappa\log\left(\frac{U_d^n}{\mu_{od}^n}\right) - \log\left(\sum_{\delta}\left(\frac{U_\delta^n}{\mu_{o\delta}^n}\right)^{\kappa}\right), \\
&= -\kappa\log\left(\mu_{od}^n\right) \underbrace{-\log\left(\sum_{\delta}\left(\frac{U_\delta^R}{\mu_{o\delta}^n}\right)^{\kappa}\right)}_{=\iota_o} + \underbrace{\kappa\log\left(U_d^n\right)}_{=\iota_d}, \\
&= \iota_o + \iota_d - \kappa\log\left(\mu_{od}^n\right), \\
&= \iota_o + \iota_d - \kappa\log\left(\rho_0^n dist_{od}^{\rho_1^n}\exp\left(-\frac{\epsilon_{od}^n}{\kappa}\right)\right), \\
&= \iota_o + \iota_d \underbrace{-\kappa\log\left(\rho_0^n\right)}_{\iota} - \kappa\log\left(dist_{od}^{\rho_1^n}\right) - \kappa\log\left(\exp\left(-\frac{\epsilon_{od}^n}{\kappa}\right)\right), \\
&= \iota + \iota_o + \iota_d - \kappa\rho_1^n\log\left(dist_{od}\right) + \epsilon_{od}^n.
\end{aligned}
$$

This migration gravity equation states that, conditional on data on migration shares $\{\eta_{od}^n\}$, geographic distances $\{dist_{od}\}$, the migration elasticity $\{\kappa\}$, and the inclusion of origin and destination fixed effects $\{\iota_o, \iota_d\}$, then the migration cost elasticities $\{\rho_1^n\}$ can be identified in the data.

Fundamental location productivity $\{\mathcal{A}_o\}$. I recover unobserved fundamental location productivities through model inversion. Given values for parameters $\{\sigma, \widetilde{\gamma}\}$, trade costs $\{\tau_{od}\}_{\forall o,d\in\mathcal{S},\mathcal{S}}$, and data on wages and population $\left\{w_o^L, w_o^R, L_o, R_o\right\}_{\forall o\in\mathcal{S}}$, there is a unique set of values for fundamental location productivities $\{\mathcal{A}_o\}_{\forall o\in\mathcal{S}}$ that is consistent with the data. Given equilibrium in goods market (32), trade shares (23), and aggregate productivity (21), I construct the following system of excess demand functions:

$$\begin{aligned}
\mathbb{D}_o(\mathcal{A}) &\equiv w_o^L L_o + w_o^R R_o - \sum_d \chi_{od}\left(w_d^L L_d + w_d^R R_d\right), \\
&= w_o^L L_o + w_o^R R_o - \sum_d \frac{A_o\left(\tau_{od} w_o^L\right)^{1-\sigma}}{\sum_o A_o\left(\tau_{od} w_o^L\right)^{1-\sigma}}\left(w_d^L L_d + w_d^R R_d\right),
\end{aligned}$$

where A_o is a function of location fundamentals $\mathcal{A}_o$. It can be shown that this excess demand functions are (i) continuous, (ii) homogeneous of degree zero, (iii) $\sum_o \mathbb{D}_o(\mathcal{A}) = 0$, and (iv) $\frac{\partial \mathbb{D}_o(\mathcal{A})}{\partial \mathcal{A}_l} > 0, \forall o, l \in \mathcal{S}, \mathcal{S}, l \neq o$ and $\frac{\partial \mathbb{D}_o(\mathcal{A})}{\partial \mathcal{A}_o} < 0, \forall o \in \mathcal{S}$. Given this properties, up to a normalization, there exists a unique vector $\mathcal{A}^*$ such that $\mathbb{D}_o(\mathcal{A}^*) = 0, \forall o \in \mathcal{S}$. I use data on wages and population $\left\{w_o^L, w_o^R, L_o, R_o\right\}_{\forall o \in \mathcal{S}}$ for year 2014.

Fundamental location amenities $\{\mathcal{B}_o^n\}$. I recover unobserved fundamental location amenities through model inversion. Given values for parameters $\{\sigma, \kappa, \widetilde{\gamma}\}$, trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, migration costs $\{\mu_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$, fundamental location productivities $\{\mathcal{A}_o\}_{\forall o \in \mathcal{S}}$, and data on wages and population $\left\{w_o^L, w_o^R, L_o, R_o\right\}_{\forall o \in \mathcal{S}}$, there is a unique set of values for fundamental location amenities $\{\mathcal{B}_o^n\}_{\forall o \in \mathcal{S}}^{n=\{L,R\}}$ that is consistent with the data. Given labor supply functions (27) and (29), migration shares (20), and indirect utility functions (18), I construct the following system of excess demand functions:

$$\begin{aligned}
\mathbb{D}_d^R\left(\mathcal{B}^R\right) &= R_d - \sum_o \eta_{od}^R R_o, \\
&= R_d - \sum_o \left(\frac{\left(\frac{U_d^R}{\mu_{od}^R}\right)^\kappa}{\sum_\delta \left(\frac{U_\delta^R}{\mu_{o\delta}^R}\right)^\kappa}\right) R_o, \\
&= R_d - \sum_o \left(\frac{\left(\frac{\mathcal{B}_d^R w_d^R}{\mu_{od}^R P_d}\right)^\kappa}{\sum_\delta \left(\frac{\mathcal{B}_\delta^R w_\delta^R}{\mu_{o\delta}^R P_\delta}\right)^\kappa}\right) R_o.
\end{aligned}$$

The same procedure can be applied for workers:

$$\mathbb{D}_d^L\left(\mathcal{B}^L\right) = L_d - \sum_o \left(\frac{\left(\frac{\mathcal{B}_d^L w_d^L}{\mu_{od}^L P_d}\right)^\kappa}{\sum_\delta \left(\frac{\mathcal{B}_\delta^L w_\delta^L}{\mu_{o\delta}^L P_\delta}\right)^\kappa}\right) L_o.$$

Prices $\{P_d\}_{\forall d \in \mathcal{S}}$ are constructed given Equations (22) and (21). It can be shown that these excess demand functions are (i) continuous, (ii) homogeneous of degree zero, (iii) $\sum_o \mathbb{D}_d^n(\mathcal{B}^n) = 0$, and (iv) $\frac{\partial \mathbb{D}_d^n(\mathcal{B}^n)}{\partial \mathcal{B}_l^n} > 0, \forall d, l \in \mathcal{S}, \mathcal{S}, l \neq o$ and $\frac{\partial \mathbb{D}_d^n(\mathcal{B}^n)}{\partial \mathcal{B}_d^n} < 0, \forall d \in \mathcal{S}$.

Given this properties, up to a normalization, there exists a unique vector $\mathcal{B}^{n^*}$ such that $\mathbb{D}_d^n\left(\mathcal{B}^{n^*}\right) = 0, \forall d \in \mathcal{S}, n = \{L, R\}$. I use data on wages and population $\left\{w_o^L, w_o^R, L_o, R_o\right\}_{\forall o \in \mathcal{S}}$ for year 2014.

G Solution algorithms

In this section I describe the algorithm that solves the model. The supra-script (i) denotes a variable as an "input", and the supra-script (o) denotes a variable as an "output".

G.1 Equilibrium

Given the exogenous distribution of workers and inventors across locations $\left\{\overline{L}_o, \overline{R}_o\right\}_{\forall o \in \mathcal{S}}$, location fundamentals $\left\{\mathcal{A}_o, \mathcal{B}_o^L, \mathcal{B}_o^R\right\}_{\forall o \in \mathcal{S}}$, iceberg migration costs $\{\mu_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$, iceberg trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, and values for all parameters of the model, the model is solved following these steps:

1. Guess $\left\{w_o^{L(i)}, w_o^{R(i)}, A_o^{(i)}\right\}_{\forall o \in \mathcal{S}}$ and $\overline{\pi}^{(i)}$:

   (a) Bilateral price indices $\{P_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$:

   $$P_{od}^{1-\sigma} = A_o^{(i)}\left(\overline{m}\tau_{od}w_o^{L(i)}\right)^{1-\sigma}$$

   (b) Price indices $\{P_d\}_{\forall d \in \mathcal{S}}$:

   $$P_d = \left(\sum_o P_{od}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$$

   (c) Migration shares $\{\eta_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$:

   $$\eta_{od}^n = \frac{\left(\frac{\mathcal{B}_d^n w_d^{n(i)}}{\mu_{od}^n P_d}\right)^\kappa}{\sum_\delta \left(\frac{\mathcal{B}_\delta^n w_\delta^{n(i)}}{\mu_{o\delta}^n P_\delta}\right)^\kappa}$$

   (d) Number of workers and inventors $\{L_d, R_d\}_{\forall d \in \mathcal{S}}$:

   $$L_d = \sum_o \eta_{od}^L \overline{L}_o,$$
   $$R_d = \sum_o \eta_{od}^R \overline{R}_o$$

(e) Income $\{Y_o\}_{\forall o\in\mathcal{S}}$:

$$Y_o = \left(1+\overline{\pi}^{(i)}\right)\left(w_o^{L^{(i)}}L_o + w_o^{R^{(i)}}R_o\right)$$

(f) Expenditure equals income:

$$X_o = Y_o$$

(g) Location profits $\{\overline{\pi}_o\}_{\forall o\in\mathcal{S}}$:

$$\overline{\pi}_o = \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma}\psi}{\sigma}\frac{\mathcal{A}_o R_o^{\widetilde{\gamma}}}{w_o^{R^{(i)\zeta}} w_o^{L^{(i)\sigma-1}}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d\right)^{\frac{1}{1-\zeta}}$$

(h) Profits per-capita $\overline{\pi}^{(o)}$:

$$\overline{\pi}^{(o)} = \frac{1}{N}\sum_o \overline{\pi}_o,$$

where $N = L + R$ is total population

(i) New worker wages $\left\{w_o^{L^{(o)}}\right\}_{\forall o\in\mathcal{S}}$:

$$w_o^{L^{(o)}} = \overline{m}^{-1}\left(\frac{A_o^{(i)}}{L_o}\sum_d \tau_{od}^{-\sigma}P_d^{\sigma-1}X_d\right)^{\frac{1}{\sigma}}$$

(j) New inventor wages $\left\{w_o^{R^{(o)}}\right\}_{\forall o\in\mathcal{S}}$:

$$w_o^{R^{(o)}} = \frac{\zeta\overline{m}^{1-\sigma}}{\sigma}\frac{\psi\mathcal{A}_o R_o^{\widetilde{\gamma}-(1-\zeta)}}{w_o^{L^{(i)\sigma-1}}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d$$

(k) Normalize wages such that $w_1^{L^{(o)}} = 1$

(l) New location productivity $\left\{A_o^{(o)}\right\}_{\forall o\in\mathcal{S}}$:

$$A_o^{(o)} = \left(\left(\frac{\zeta\overline{m}^{1-\sigma}}{\sigma}\right)^{\zeta}\left(\psi\mathcal{A}_o R_o^{\widetilde{\gamma}}\right)\left(w_o^{L^{(i)\sigma-1}}w_o^{R^{(i)}}\right)^{-\zeta}\left(\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d\right)^{\zeta}\right)^{\frac{1}{1-\zeta}}$$

(m) Update $w_o^{R^{(i)}} = w_o^{R^{(o)}}$, $w_o^{L^{(i)}} = w_o^{L^{(o)}}$, $A_o^{(i)} = A_o^{(o)}$ and $\overline{\pi}^{(i)} = \overline{\pi}^{(o)}$

(n) Iterate until convergence is achieved

G.2 Equilibrium with R&D subsidies

Given the exogenous distribution of workers and inventors across locations $\left\{\overline{L}_o, \overline{R}_o\right\}_{\forall o \in \mathcal{S}}$, location fundamentals $\left\{\mathcal{A}_o, \mathcal{B}_o^L, \mathcal{B}_o^R\right\}_{\forall o \in \mathcal{S}}$, iceberg migration costs $\{\mu_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$, iceberg trade costs $\{\tau_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$, R&D subsidies $\{s_o\}_{\forall o \in \mathcal{S}}$, and values for all parameters of the model, the model is solved following these steps:

1. Guess $\left\{w_o^{L^{(i)}}, w_o^{R^{(i)}}, A_o^{(i)}\right\}_{\forall o \in \mathcal{S}}$ and $\left\{\overline{\pi}^{(i)}, \tau^{(i)}\right\}$:

   (a) Bilateral price indices $\{P_{od}\}_{\forall o,d \in \mathcal{S},\mathcal{S}}$:

   $$P_{od}^{1-\sigma} = A_o^{(i)} \left(\overline{m}\tau_{od} w_o^{L^{(i)}}\right)^{1-\sigma}$$

   (b) Price indices $\{P_d\}_{\forall d \in \mathcal{S}}$:

   $$P_d = \left(\sum_o P_{od}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$$

   (c) Migration shares $\{\eta_{od}^n\}_{\forall o,d \in \mathcal{S},\mathcal{S}}^{n=\{L,R\}}$:

   $$\eta_{od}^n = \frac{\left(\frac{\mathcal{B}_d^n w_d^{n^{(i)}}}{\mu_{od}^n P_d}\right)^{\kappa}}{\sum_\delta \left(\frac{\mathcal{B}_\delta^n w_\delta^{n^{(i)}}}{\mu_{o\delta}^n P_\delta}\right)^{\kappa}}$$

   (d) Number of workers and inventors $\{L_d, R_d\}_{\forall d \in \mathcal{S}}$:

   $$L_d = \sum_o \eta_{od}^L \overline{L}_o,$$
   $$R_d = \sum_o \eta_{od}^R \overline{R}_o$$

   (e) Income $\{Y_o\}_{\forall o \in \mathcal{S}}$:

   $$Y_o = \left(1 + \overline{\pi}^{(i)} + \tau^{(i)}\right)\left(w_o^{L^{(i)}} L_o + w_o^{R^{(i)}} R_o\right)$$

   (f) Expenditure equals income:

   $$X_o = Y_o$$

(g) Location profits $\{\overline{\pi}_o\}_{\forall o\in\mathcal{S}}$:

$$\overline{\pi}_o = \left(\frac{\kappa_\zeta \overline{m}^{1-\sigma}\psi}{\sigma}\frac{\mathcal{A}_o R_o^{\widetilde{\gamma}}}{(1-s_o)^\zeta\, w_o^{R(i)\zeta}\, w_o^{L(i)\sigma-1}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d\right)^{\frac{1}{1-\zeta}}$$

(h) Profits per-capita $\overline{\pi}^{(o)}$:

$$\overline{\pi}^{(o)} = \frac{1}{N}\sum_o \overline{\pi}_o,$$

where $N = L + R$ is total population

(i) New value of tax rate $\tau^{(o)}$:

$$\tau^{(o)} = \frac{\sum_o s_o\left(w_o^{R(i)}R_o\right)}{\sum_o\left(w_o^{L(i)}L_o + w_o^{R(i)}R_o\right)}$$

(j) New worker wages $\left\{w_o^{L(o)}\right\}_{\forall o\in\mathcal{S}}$:

$$w_o^{L(o)} = \overline{m}^{-1}\left(\frac{A_o^{(i)}}{L_o}\sum_d \tau_{od}^{-\sigma}P_d^{\sigma-1}X_d\right)^{\frac{1}{\sigma}}$$

(k) New inventor wages $\left\{w_o^{R(o)}\right\}_{\forall o\in\mathcal{S}}$:

$$w_o^{R(o)} = \frac{\zeta\overline{m}^{1-\sigma}}{\sigma}\frac{\psi\mathcal{A}_o R_o^{\widetilde{\gamma}-(1-\zeta)}}{(1-s_o)\, w_o^{L(i)\sigma-1}}\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d$$

(l) Normalize wages such that $w_1^{L(o)} = 1$

(m) New location productivity $\left\{A_o^{(o)}\right\}_{\forall o\in\mathcal{S}}$:

$$A_o^{(o)} = \left(\left(\frac{\zeta\overline{m}^{1-\sigma}}{\sigma}\right)^\zeta\left(\psi\mathcal{A}_o R_o^{\widetilde{\gamma}}\right)\left((1-s_o)\, w_o^{L(i)\sigma-1}\, w_o^{R(i)}\right)^{-\zeta}\left(\sum_d \tau_{od}^{1-\sigma}P_d^{\sigma-1}X_d\right)^\zeta\right)^{\frac{1}{1-\zeta}}$$

(n) Update $w_o^{R(i)} = w_o^{R(o)}$, $w_o^{L(i)} = w_o^{L(o)}$, $A_o^{(i)} = A_o^{(o)}$, $\overline{\pi}^{(i)} = \overline{\pi}^{(o)}$, and $\tau^{(o)} = \tau^{(i)}$

(o) Iterate until convergence is achieved