THE UNIVERSITY OF CHICAGO

ESSAYS IN MACROECONOMICS AND ECONOMIC GROWTH

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

BY
IHSAN FURKAN KILIC

CHICAGO, ILLINOIS
AUGUST 2024

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my committee members for their invaluable guidance and support. I am especially indebted to my chair, Ufuk Akcigit, for his outstanding mentorship, insightful advice, and continuous encouragement throughout the program. His expertise and dedication have been instrumental in shaping my research. I also extend my sincere appreciation to Mikhail Golosov and Veronica Guerrieri for their guidance and steadfast support.

I would like to extend my heartfelt thanks to Craig Chikis and Jun Wong for their invaluable collaboration in the early stages of the first chapter. I am also deeply grateful to my coauthors of the second chapter, Ufuk Akcigit, Sina Ates, Matthias Mertens, and Steffen Mueller, for their collaborative efforts and commitment to our research. I appreciate the insightful comments from Ibrahim Bicak, Sunduz Divle, Shanon Hsuan-Ming Hsu, Jeremy Pearce, Berkay Saygin, and Younghun Shim, as well as seminar participants and ART group members in Chicago and Germany.

I would like to express my most sincere and heartfelt gratitude to my family. To my mother, Saadet Kilic, for her unwavering support and encouragement throughout my entire life. Her dedication to my education and her constant belief in my potential have been the driving forces behind my achievements. To my brother, Etka Kilic, for his steadfast support and the deep companionship we share, which has been a source of strength and comfort. To my late father, Ismail Kilic, whose memory continues to inspire me. And to my beloved wife, Sena Aytis Kilic, for her endless love, patience, and unwavering belief in me. Her steadfast support and devotion have been my guiding light, and I am forever grateful for her presence, which has made every step of this journey more meaningful and joyous.

ABSTRACT

This dissertation encompasses two chapters.

The first chapter examines the spatial distribution of inventors in the U.S. economy and its implications for aggregate innovation and economic growth. It highlights the pronounced geographical concentration of innovative activity within select U.S. innovation hubs. Utilizing patent citation data, significant heterogeneity and imperfections in knowledge spillovers across states are identified. In light of these findings, a novel endogenous growth model is developed, featuring mobile inventors and workers between states. The model is equipped with an exogenous knowledge network that captures the dynamics of idea exchange across locations. In the model, inventors do not internalize the effect of their location choice on the diffusion of ideas to the rest of the economy, necessitating a place-based R&D subsidy policy to maximize the potential of knowledge linkages. Location specific amenities and exogenous research productivities are recovered by matching observed inventor and worker allocations in space. The knowledge diffusion network is estimated from patent citation data, and optimal policy is analyzed by a set of counterfactual exercises. The policy implies a greater concentration of inventors in established innovation hubs, enhancing welfare by 1.8% in consumption equivalent terms and elevating the economy's growth rate by 0.14 percentage points.

The second chapter examines the slow economic convergence of East Germany after it is unified with West Germany. Even 30 years after the reunification, regions in East Germany (the former socialist GDR) live in considerably different economic conditions, with the average GDP per capita still about 20 percent below the average level in the West German regions. In this paper, we explore the obstacles that impeded full convergence despite massive support to the East with a particular focus on technological differences and firm behavior. In the immediate aftermath of the reunification, production in the former GDR exhibited a rapid catch-up with the West with a pick-up in labor productivity. But the convergence then lost steam quickly with a stark difference between East and West German firms' product qualities persisting ever since. We build a quantitative model of innovation, competition, and regional integration that is able to replicate these dynamics and provides a suitable setting to evaluate alternative policies that could have altered these dynamics. We find that delaying the reunification—i.e., opening up to competition from the West—would not help the Eastern firms to build up capacity. Sustained support for R&D in the East from the West could have helped shrink persistent gaps in product quality and income, although more effective alternatives appear to be subsidies to Western firms via either R&D support, with knowledge spillovers lifting also Eastern technology, or direct income support to facilitate technology transfer to the East via licensing.

CHAPTER 1

SPATIAL ALLOCATION OF INVENTORS IN THE US AND ECONOMIC GROWTH

1.1 Introduction

Silicon Valley in California, Route 128 in Massachusetts, and Seattle in Washington are some of the most important innovation hubs in the US. In 2005, top 10 US states in terms of patenting produced 65% of all the patents in the US. Similarly, these states were home to 64% of all inventors in the same year. Concentration of innovative activity is pervasive, and pointed out in several studies (Moretti [2021], Carlino and Kerr [2015] for a survey of the literature), but this observation is not only specific to the US economy. For example, Carrincazeaux et al. [2001] shows that six regions in France account for 75% of corporate R&D workers while their share of production workers is only 45%.

Spatial allocation of inventors (or researchers and scientists in general) is crucial for the overall growth trajectory of an economy. Main reason can be traced back to Marshall [2009] in his treatment that localities benefit from knowledge spillovers within locations. Types of spillovers that are important for the creation of new inventions are in intertemporal nature, and can be traced by patent citations just like academic citations. Inventors build on top of the shoulders of the past when they develop new ideas. If knowledge diffusions between innovation centers are not frictionless, then inventor's location choices, or the accumulation of innovative activity in certain regions would have nontrivial consequences on the technological development of the overall economy.

In this paper, I study intertemporal knowledge spillovers and spatial allocation of inventors across locations in the United States, and its consequences on the growth of the US economy. Firstly, I document signification variation in the spatial allocation of inventors relative to workers across US states. Inventors relocate intensely to a few number of states that are usually home to innovation hubs. What is the effect of their location choice on the innovation process and idea creation? In the literature, researchers approach to this question in terms of knowledge spillovers (Moretti [2021]) and agglomeration economies. In this paper, I take a more general approach by extending the spatial aspect of intertemporal knowledge spillovers. That is, rather than focusing only on within-location agglomeration externalities, I study knowledge flows across all regions, which are proxied by patent citations. Patent citations provide many advantages in this regard, as Jaffe et al. [1993] and Jaffe et al. [2000] argue, they can be interpreted as paper trails of knowledge spillovers between inventors, although inventors might live very far away from each other in space.

Knowledge spillovers, in the specific context of innovation and R&D, are mostly in intertemporal nature, finding its meaning in the famous phrase, new ideas are build on top of the shoulders of the past giants. Current inventors build on past inventions when they create new ideas, and they cite the ideas from which they benefit the most (Jaffe et al. [2000]). Analysis of patent citations data reveals that patent citations are also spatially concentrated. In particular, states like California, Massachusetts, Connecticut are the most cited states in the US. However, I argue that this observation alone cannot be interpreted as the importance of these states being the sole origins of idea creation hubs in the US economy. Instead, I investigate patent citation lags between locations, and find significant variation across state-pairs, which suggests that some states are better connected to each other in the sense that they tend to cite each other relatively quicker than other states. This variation in patent citation lags identifies the extent of connections between locations in the US in terms of knowledge flows.

I study knowledge spillovers because of several reasons. In endogenous growth literature, intertemporal knowledge spillovers are identified as the main reason for the justification of R&D subsidy policies (Aghion and Howitt [1992]). When inventors invent their ideas and develop new techniques, they do not internalize their effect on the creation of future ideas. This positive externality creates an incentive for the policy maker in favor of R&D subsidies. However, if knowledge spillovers are heterogeneous across location-pairs, and if they are imperfect, then is there a room for place-based R&D subsidy policies rather than a homogeneous one across space, and if so, how should it be designed optimally?

Answering this normative question requires a theory of intertemporal knowledge spillovers between locations, inventor migration, and innovation as the engine of economic growth. I build a novel spatial economic growth model with endogenous inventor and worker migration choices, and knowledge linkages across locations. In the model, inventors and workers who hold idiosyncratic preferences for locations can move between regions freely albeit subject to a simple timing friction. Their objective is to maximize their life-time utility. Worker side of the model is intentionally kept simple. In equilibrium, workers earn the same wage income in all locations, thus their migration decision identifies location specific characteristics that are also common to inventors when they relocate. These characteristics are called amenities in the model. On top of amenity differences, locations in the model are heterogeneous in their fundamental research productivities. The idea is that some locations provide more resources for the R&D process such as presence of leading universities like Stanford in California, or certain institutions supporting innovation and dynamism. These characteristics are unobserved to the researcher, and they are estimated by matching observed inventor allocation across US states. Another dimension by which locations are differentiated is their connectedness to the rest of the economy in terms of idea flows. In the model, past ideas invented in locations spill over to the rest over time, and they form the endogenous idea stock over which inventors in destination locations build on top when they perform R&D. The diffusion process is subject to frictions in that past ideas diffuse to other regions with a random time lag, where average time lag is specific to the location-pair, and it does not have to be symmetric. Inventors benefit both from location's fundamental (exogenous) resources and endogenous idea accumulation. All else equal, inventors that are located in states to which idea inflows are faster would be more productive in R&D, and they would earn higher wages. In the model, this is the source of externality of inventor location choice on the rest of the economy.

The model also addresses endogenous location choices of firms who are employers of inventors. Potential entrants in the model are mobile across locations, and in equilibrium, they move to the location that provide the highest discounted future profits. Firms that enter to the market has a single R&D lab in a location in which inventors are employed, while they are indifferent across locations for the place of their production. In equilibrium, more research productive locations are home to a higher number of innovative firms who demand more for inventors. Thus, the model simultaneously explains the presence of high volume of innovative firms in innovation hubs along with high numbers of inventors relocated there.

I quantify the model to recover location specific amenities and exogenous research productivities from the observed allocation of production workers and inventors. The knowledge network, which is the matrix of diffusion rates across state-pairs are estimated from patent citation lags in line with the idea accumulation process in the model. Estimation of the network reveals that there is heterogeneity in connectedness between locations, within-location spillovers are the strongest, and estimated idea diffusion rates increase with physical proximity of states, and academic citation flows between states. After estimating the parameters, I test the model fit for a set of untargeted moments. This exercise suggests that the model performs well in explaining these moments.

I report my estimation results in two steps. In the first step, I estimate the model assuming that that the US economy is comprised of only ten states in which most of the patents in the US are produced. I compare estimation results for two versions of the model—with and without knowledge spillovers. Then, I proceed with the optimal policy and run several counterfactual exercises to understand the nature of the policy. The optimal policy calls for more concentration in inventor allocation in the space while allocating inventors to more central states in the knowledge network. In the second step, I estimate the model for all the US states. The optimal place based R&D subsidy policy concentrates inventors mostly in Washington, California and Massachusetts, although the model is abstracted away from reduced form agglomeration spillovers. The reason is that these states are connected well with the rest of the economy, while they are also the most research productive states. The welfare of the society increases by 1.8% in consumption equivalent terms under the optimal policy. Most of the increase in welfare stems from increased growth rate of the economy with a 0.14 percentage points.

1.2 Empirical regularities

In this section, I focus on two strong phenomena that I observe from the patent data. Firstly, I show that innovation and patenting, and relatedly inventor locations are clustered geographically in the United States, more than other types of economic activity such as employment, population, and GDP. Then, I proceed with the distribution of inventors per employment in locations to understand the extent of the concentration. Some states are asymmetrically populated by inventors relative to overall employment, suggesting heterogeneous demand for researchers across US states. Finally, I analyze geographical concentration in patent citations. I document huge variation in patent citation lags between state-pairs as an evidence for varying degrees of knowledge linkages across US states.

Data. The period I study in this paper is the decade between 2000 and 2010. The geographical unit of analysis is chosen to be all the states in the US including DC (51 states). Worker and inventor allocations over the cross section of states are measured for the median year of the analysis, 2005. For patent citations data, I restrict citing patents between 2000-2010, and cited patents between 1990-2010 (10 more years before 2000). The reason I restrict the sample at 2010 is to remove truncation bias in patent citations as suggested by Hall et al. [2001]. The other reason not to include cited patents that are issued before 1990 is to ensure computational feasibility in my estimation procedure.

State level employment data comes from Census Bureau's Business Dynamics Statistics (BDS). In BDS data, inventors and researchers (in general any type of employment associated with R&D and innovation) are also counted under total employment figures. As the number of total researchers in the US is relatively very small compared to total employment, I do not subtract number of inventors from state level total employment numbers.[1]

For patent citations and spatial allocation of inventors, I use PatentView's disambiguated patent dataset. This dataset is a good fit for the purpose of this study as inventor names and addresses are disambiguated, and hence, allocation of inventors across US states is measurable. The dataset consists of the universe of patents that are applied to USPTO, and it covers the years starting from 1976 until recent years. As explained above, the time range for patent citations is restricted to 2000-2010 for citing patents, and 1990-2010 for cited patents. Hence the maximum citation lag in my data is 20 years with the shortest being zero—citing and cited patents belong to the same year. To determine the location of patents, I use inventor addresses assuming that the R&D is performed in the same location where inventor lives, in line with the model as will be explained in Section 1.3. Finally, to determine the spatial allocation of inventors across US states, I count the number of unique inventors in 2005, similar to employment.

It should be noted that in the patent data, only inventors that apply for a patent in a year are observed. Therefore, interpreting the number of inventors in the patent dataset as a direct correspondence of the total number of inventors/researchers in states, including ones that do not apply for a patent, is biased, and it is corrected through the lens of the model when I estimate the model parameters, as the model results in several predictions on

1. This is more of a practical approach as it simplifies the model inversion procedure by removing one iterative step.

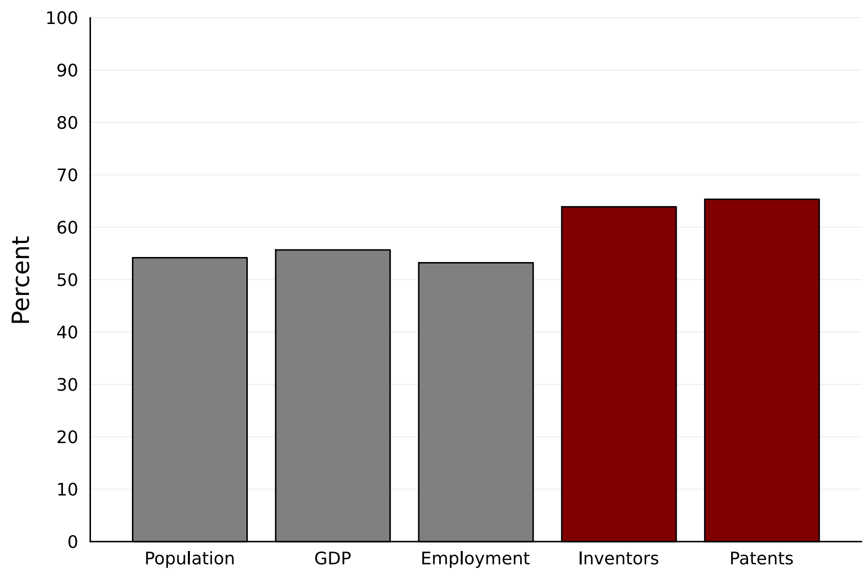


Figure 1.1: Share of top 10 states with respect to type of economic activity

Note: The set of top 10 states changes among five activities.

patenting probability of individual inventors. Full details are described in Section 1.4.

1.2.1 Spatial allocation of inventors and employment

Innovative activity that is measured by inventors and patenting is spatially more concentrated than other types of economic activities such as population, GDP, and employment. Figure 1.1 shows the higher concentration in inventor and patenting allocation across US states as measured by the share of top 10 states.

Figure 1.2 shows the scatter plot of inventor and employment shares of all US states. California is home to the highest share of employment and inventors in the US. There is strong positive correlation between states' inventor and employment shares. This figure also shows that inventor share of some states such as California (CA), Washington (WA) and Massachusetts (MA) is higher than their employment share. On the contrary, small states such as Alaska (AK), Wyoming (WY), and South Dakota (SD) are mostly populated with employees rather than inventors.

Figure 1.3 visualizes the variation in inventor-to-employment ratio on a US map. In-

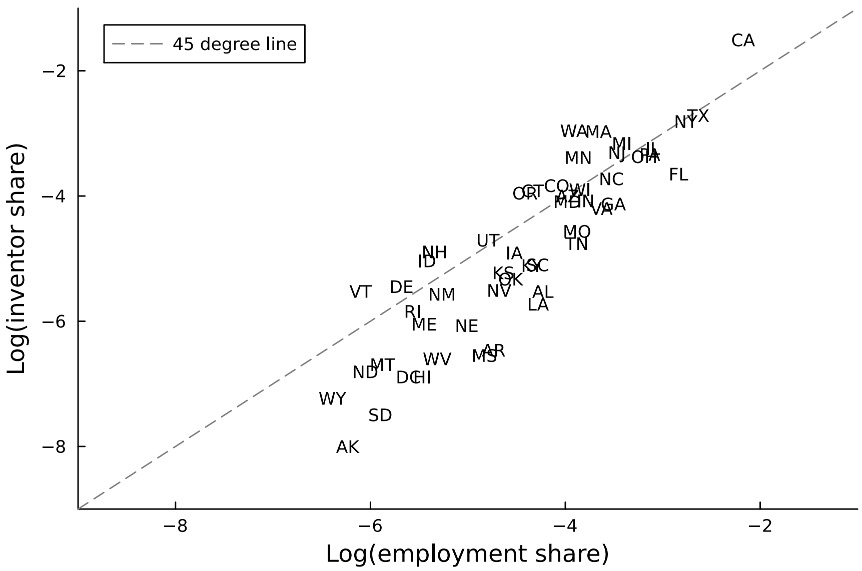


Figure 1.2: Employment and inventor shares of US states, 2005

Note: Axes are in log scale for visibility. Worker share of a state is calculated as the ratio of total employment of the state divided by the total employment in the US. Similarly, inventor share of a state equals to the ratio of number of inventors located in the state to the total number of inventors.

ventors prefer some clusters of states more than workers. For example, in the West coast, close states such as Washington, Oregon, California and Idaho are populated by inventors more than workers. Washington is the highest state in terms of inventor to worker ratio with more than three inventors per thousand employment. In the Midwest, Minnesota and Michigan stand out as the states that are preferred more by inventors. In the East cost, on the other hand, Massachusetts, Connecticut, Vermont and New Hampshire form a cluster where inventors are relocated relatively more than workers.

States to which inventors are relocated more intensely overlaps partially with the set of states that produce most of the patents in the US in year 2005. These states account for 65% of patents, and 64% of all inventors in the US. For instance, Washington, California, Minnesota, Michigan, and Massachusetts from above are also in the list of top 10 patenting states. Other states in the top 10 list are Texas, New Jersey, New York, Illinois and Pennsylvania. However, inventors do not relocate towards these states disproportionately. This observation suggests that a state's share of inventors alone might not be very informative

Number of inventors per thousand employment

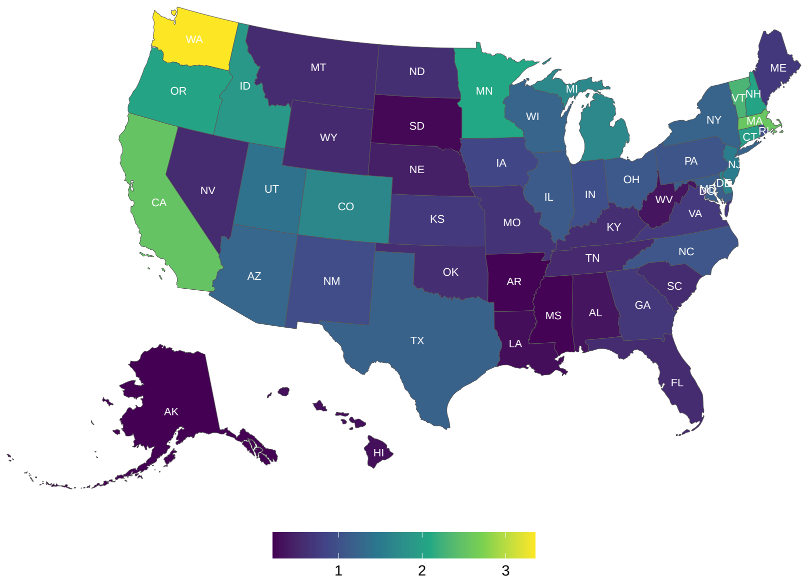


Figure 1.3: Number of inventors per thousand employment, 2005

about its research productivity and the extent of innovative activity, as there are other reasons that could explain the high number of inventors in a state such as amenities, which would also affect the migration choice of production workers in a similar way. Therefore, the variation in inventor-to-worker ratio across locations is a better candidate to identify the specific factors that applies only to inventors when they decide where to locate.

The line of reasoning above is an articulation of a simple supply-demand analysis. Supply of inventors to a region increases with the wage rate offered there. Amenities on the other hand can be considered as a supply shifter. The demand for inventors in a region decreases in inventor wages, while research productivity of the state can be considered as a demand shifter. Higher number of inventors in a state can be the equilibrium outcome of both supply and demand factors. High amenity states would attract more inventors all else equal. However, amenities would also affect the migration decision of workers. Therefore, after controlling for state's employment share, the remaining variation in observed inventor shares stems from demand related factors such as research productivity of the location.

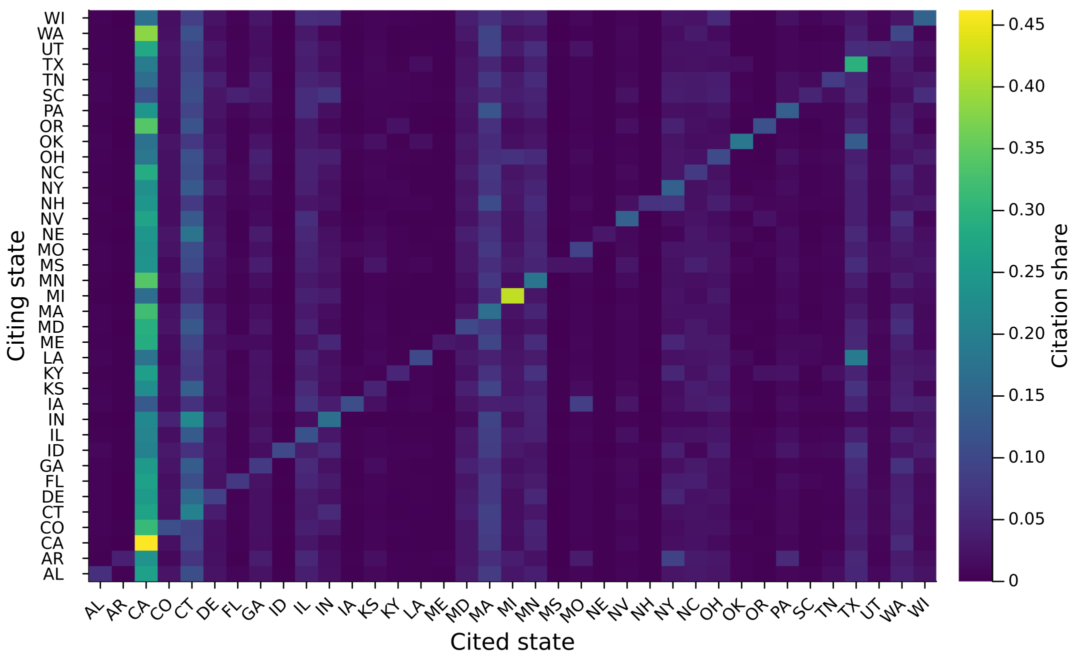


Figure 1.4: Patent citation shares across US states

1.2.2 Patent citations

In the sample patent dataset, there are 1,870,743 unique patents that applied for grant between the years 1990 and 2010. Among these, 1,029,211 of them are applied after 2000. The number of observed citations from patents between 2000-2010 to patents from 1990-2010 is 15,727,544.

Figure 1.4 plots the heatmap of citation shares of citing-cited state pairs between years 2001 and 2010. In particular, the columns correspond to cited states, while citing states are represented by rows. A cell that corresponds to a citing state A and a cited state B shows the share of citations received by patents from B in total citations given by A's patents. As inferred from the citation share matrix, within citation rates are usually higher than out-of-state citation shares. That is, patents are more likely to cite other patents that originate from their states. Furthermore, some cited states such as California (CA), Connecticut (CT), Massachusetts (MA), Texas (TX) stand out as the most cited states by others. It is not a coincidence that these states capture a higher fraction of citations from other states, as they are also the states that produce most of the patents in the US. The citation share of small states is expected to be lower, as the number of patents available for citation in these states is limited. Therefore, Figure 1.4 is not very informative about the flow of knowledge across states when taken at its face value. However, it still shows, just like patenting activity in the US, patent citations are also concentrated towards certain states.

In this study, I propose another measure for idea flows between locations from the patent citation data—heterogeneous citation lags between states measured as the time lag between citing and cited patents. This moment is more informative about the pace at which ideas flow across states. Figure 1.5 shows an example in this regard. For visual clarity, I pick the top 6 states in terms of patenting; California (CA), Washington (WA), Texas (TX), New York (NY), Michigan (MI). Each of six plots represent a citing state. Lines correspond to cited states excluding the citing state itself in order to focus on citation lags between locations. The y-axis plots the average citation probability (multiplied by 1000) calculated as the number of citations divided by the product of the number of patents in citing and cited states (this product gives the number of all possible bilateral connections). The average is taken across citing and cited years for a particular time lag. For example, the citation probability for the lag 5 (years) is the average of citation probabilities that are observed between any two years that has a lag of 5 years such as citing patents are issued in 2001 and cited patents are issued in 1996, citing patents are issued in 2002 and cited patents are issued in 1997, and so on. Thus, Figure 1.5 illustrates citation probability between any two states as a function of time lag between the time of citation and the time of creation of cited idea.

There are three general messages of Figure 1.5. The first one is that citation probability is very close to zero when the time lag between citing and cited patents is very short. In other words, the most recent patents receive very little citations. The second message is

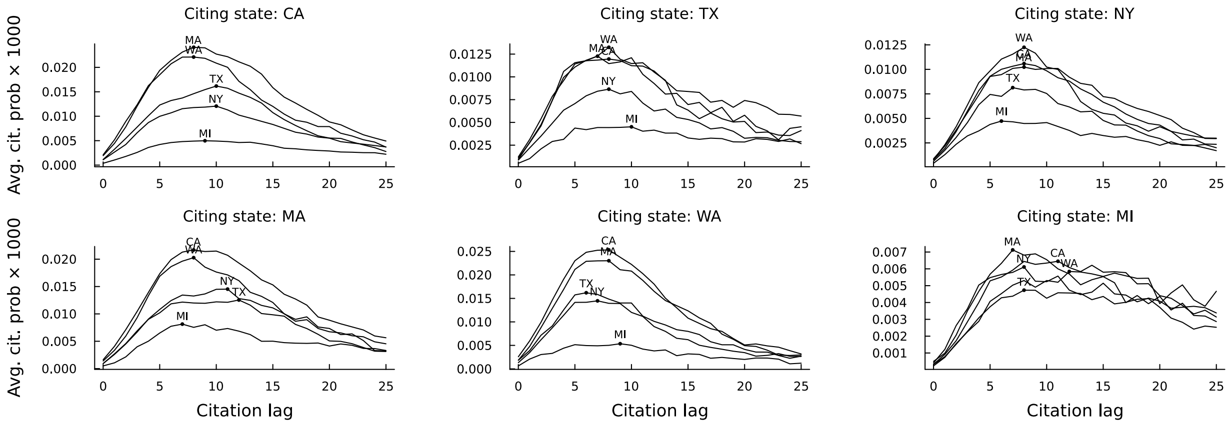


Figure 1.5: Citation probability of top 6 states as a function of time lag

Note: Within state citations are excluded from figures. For example, in the top left figure, the line corresponding to CA is excluded. Similarly, in the middle top figure, the line corresponding to TX is excluded. The reason is that within-state citation probabilities are much higher than others in levels shadowing the variation across states.

that citation probability declines as the time lag between citing and cited patents become very large. In other words, old patents are cited less frequently. Finally, citation probability peaks at moderate lags. In other words, certain amount of time is needed for cited patents to be known to others in order to start receiving citations.

Figure 1.5 reveals another interesting heterogeneity between citing-cited state pairs, i.e. the time lag at which citation probability hits the peak varies across states. For instance, let's focus on the top left figure in which the citing state is California. The probability that patents from Massachusetts (MA) gets citations from patents in California (CA) peaks around a time lag of 7-8 years, while this probability peaks around 10 years for cited Texas (TX) patents. It suggests, on average, CA patents cite MA's patents earlier than TX's patents. Similarly, top middle figure, in which the citing state is Texas (TX), suggests that TX's patents cite Massachusetts's patents (MA) earlier than Michigan's patents (MI). Peak citation lags are not symmetric across citing-cited state pairs. An example is given by the bottom left figure in which the citing state is MA. Although TX patents cite MA patents relatively quicker, MA patents do not cite TX patents that faster. The peak citation lag for TX-MA (citing-cited) pair is around 7 years, while MA-TX peak lag is around 12 years.

The variation in citation lags between patents based on their locations identifies the idea

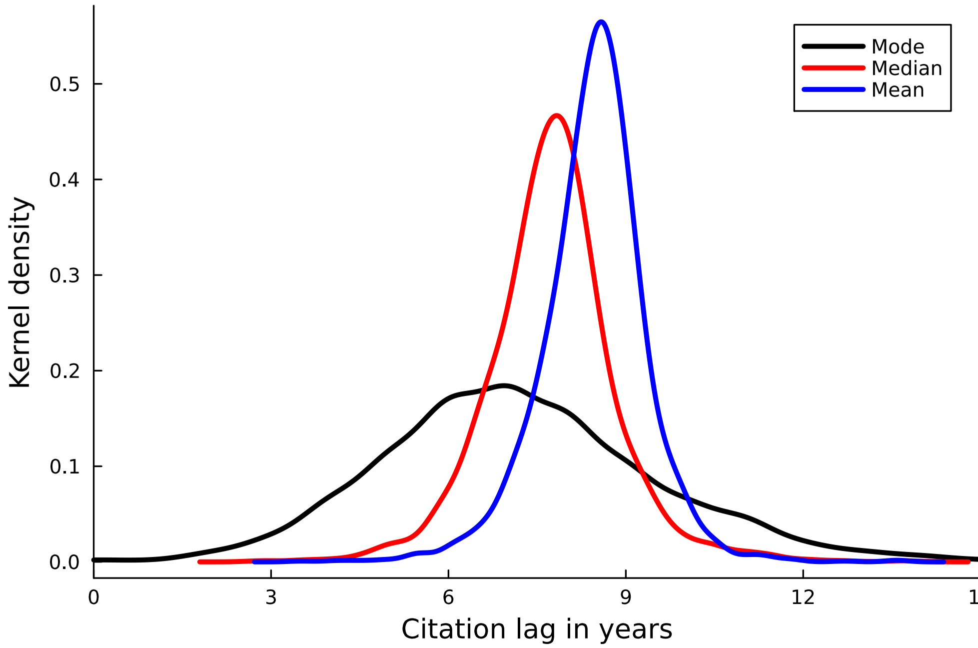


Figure 1.6: Kernel density estimates of mode, median, and mean citation lag distributions over citing-cited state pairs

Note: Each line corresponds to the distribution of a particular moment of citation lags across citing-cited state pairs. There are $51 \times 51 = 2601$ state pairs in total. For a fixed citing-cited state pair, I calculate the mode citation lag taking total citation counts at each lag being the frequency of lag observations. Then the kernel density estimate of the distribution of mode citation lag across all state pairs is plotted. The same procedure is followed for median and mean citation lag distributions.

diffusion rates across states. As suggested by the previous discussion, the mode of citation lag between any state-pair is a particular informative moment. In addition to mode citation lag, one can also consider median and mean lags being other moments that identify the average time lag in diffusion of ideas between locations. There is enough variation between state-pairs in terms of all these moments. Figure 1.6 shows the kernel density estimates of mode, median, and mean citation lag distributions over state-pairs.

Figure 1.6 reveals the extent of heterogeneity in citation lag moments across state pairs. On average, mode citation lag is shorter than median and mean citation lags suggesting a right-skewed distribution of citation lags for any given citing-cited state pair.

As will be described in Sections 1.3 and 1.4, the knowledge spillover network between states is parameterized by a matrix of idea diffusion rates. Denoting by j the state where the idea originates, and by i the state where the idea diffuses to, ω_{ij} will be defined as the (idea) diffusion rate parameter from j to i. It is assumed that ideas diffuse from j to i with a random time lag which is distributed as Exponential(ω_{ij}). Thus the average time lag in idea diffusion is given by $1/\omega_{ij}$. If ω_{ij} is higher, then ideas diffuse relatively faster from j to i helping inventors in i to benefit the past ideas of j relatively earlier. It is also assumed that ideas become obsolete with an exogenous rate of δ. As details are delegated to Section 1.4.1, the probability by which an idea originated in j will be visible in i after τ years is given by $e^{-\delta\tau}\left[1-e^{-\omega_{ij}\tau}\right]$. Assuming that ideas are cited at the time of their diffusion (conditional on not being obsolete by that time), the unknown parameter ω_{ij} will be identified from the relationship between citation probabilities and citation lags for the state pair $i \times j$. As ω_{ij} increases, inventors in i cites patents from j in earlier lags. This is the source of variation used to identify the knowledge spillover network in the US economy.

1.3 Model

Motivated by the empirical observations, I build an endogenous growth model with migration choices of two types of agents; workers and inventors. I keep the model as simple as possible by modifying important aspects. Spatial aspect of the model arises due to the fact that both workers and inventors are mobile across locations. Firms are identified by their location choices for their single R&D lab in which they hire inventors and generate new ideas for production. Simple structure of the model implies that firms are indifferent where to produce their products, however, they maximize their value—discounted sum of future profits—by choosing the location for their R&D lab at the time of entry.

1.3.1 Locations

There are N locations denoted by $i, j = 1, \ldots, N$. Locations are heterogeneous ex-ante only in two respects. Firstly, amenities provided by the location are heterogeneous and time invariant. These amenities can be considered as environmental factors, crime rates, public goods in a location which are valued identically by all individuals in the model, and are denoted by A_i. Secondly, locations differ from each other in terms of the resources they provide specifically to the inventors. These factors are, but not limited to, universities, institutions, culture, and more importantly the past ideas that have diffused to the location from the rest, helping inventors build on when they research for new innovative ideas. I denote (endogenous) research productivity of location i by $\alpha_i(t)$. This aspect of locations might be time varying as it is assumed that it is a combination of two factors. The first factor is called fundamental (exogenous) research productivity of the location, and denoted by $\bar{\alpha}_i$. The second factor is the number of past innovative ideas that have accumulated in location i, and denoted by $K_i(t)$. Researchers in location i are assumed to build on these ideas when they come up with new ideas. This respect of the model reflects the idea of intertemporal spillovers in idea generation process, which is a prominent feature of endogenous growth models, by supplementing it with a spatial aspect. More details on idea diffusion and accumulation process are described in Section 1.3.6.

I assume a specific functional form for the formation of $\alpha_i(t)$ from these factors

$$\alpha_i(t) = \bar{\alpha}_i^{1-\varphi} K_i(t)^{\varphi}, \qquad \varphi \in [0, 1] \tag{1.1}$$

That is, endogenous research productivity is concave in the number of ideas diffused, $K_i(t)$.

1.3.2 Preferences

Total population consists of two types of individuals, workers and researchers. The aggregate measure of workers in the model, denoted by $\bar{L}$, is normalized to one. Similarly, aggregate measure of inventors is also a constant parameter denoted by $\bar{R}$. Both types of individuals are mobile across locations, earn labor income, and consume hand-to-mouth every period. Specifically, they have no access to financial markets. Workers, denoted by superscript

$T = L$, work for the final good production in locations. On the other hand, researchers, denoted by superscript $T = R$, work for intermediate good firms in order to perform research and innovation. Wage rate of type T in location i is denoted by $W_i^T(t)$. Both types of agents discount future utility with a rate of $\rho > 0$.

Per period utility flow rate is denoted by $U_i^T(t)$ and equals to

$$U_i^T(t) = A_i \varepsilon_i C_i^T(t) \tag{1.2}$$

where A_i is location specific, time invariant amenity of location i, $C_i^T(t)$ is the rate of consumption of type T agent in location i. Finally, ε_i denotes individual's idiosyncratic taste for the location. Budget constraint is given by

$$C_i^T(t) = \left[1 + d(t) - \tau(t)\right] W_i^T(t) \tag{1.3}$$

For simplicity, I assume that firms are owned by a national portfolio, and profits are rebated to individuals proportional to their labor income. Therefore, $d(t) \times W_i^T(t)$ is the amount of dividends distributed to an agent with a labor income of $W_i^T(t)$. Labor income tax rate is denoted by $\tau(t)$. Tax rate is not location specific, but it can be potentially time varying, as the tax revenue collected is only used to finance place-based R&D subsidies, as explained later. Government adjusts the tax rate every period to finance its total R&D subsidy expenditures.

1.3.3 Migration

Agents update their location preferences, $\epsilon \equiv \{\varepsilon_i\}_{i=1}^N$, with a Poisson arrival rate of $\zeta > 0$. Arrival of this shock is independent across individuals, time, and locations. Location taste for each i is independently drawn from a Fréchet distribution

$$\varepsilon_i \overset{\text{ind}}{\sim} \text{Fréchet}\,(\xi, 1)$$

whose shape parameter is given by $\xi > 1$. Lower ξ indicates greater heterogeneity in location taste across individuals. After receiving new location tastes, including for the location they currently live in, individuals decide whether to move to another location or not. Migration between any pair of locations is costless.

Let $\mathcal{U}_i^T(\varepsilon, t)$ denote the discounted life-time utility, simply value, of type T living in location i with a taste of ε. Agents choose to migrate to the location that provides the highest value for themselves. That is, ex-post, they solve the following maximization problem after the realization of a location taste vector $\mathbf{e} \equiv \{e_i\}_{i=1}^N$ where e_j denotes the j$^{\text{th}}$ component of vector $\mathbf{e}$

$$\max_{j=1,\ldots,N} \mathcal{U}_j^T(e_j, t) \tag{1.4}$$

Ex-ante, location tastes are uncertain for agents. It is useful to define the following function which can be defined as the expected value of arrival of taste shocks

$$\bar{\mathcal{U}}^T(t) \equiv \int \left\{ \max_{j=1,\ldots,N} \mathcal{U}_j^T(e_j, t) \right\} f_\epsilon(\mathbf{e})\, d\mathbf{e} \tag{1.5}$$

where $f_\epsilon(\cdot)$ denotes the joint density of location taste vector.[2] The idea behind (1.5) is that agents decide to relocate to the maximum value location. Equipped with this notation, we can derive the HJB equation for the evolution of $\mathcal{U}_i^T(\varepsilon, t)$ as follows

$$\rho\mathcal{U}_i^T(\varepsilon, t) = A_i \varepsilon C_i^T(t) + \zeta\left[\bar{\mathcal{U}}^T(t) - \mathcal{U}_i^T(\varepsilon, t)\right] + \frac{\partial \mathcal{U}_i^T(\varepsilon, t)}{\partial t} \tag{1.6}$$

2. The function $\bar{\mathcal{U}}^T(t)$ does not possess a location subscript as $\mathcal{U}_j^T(e_j, t)$ is maximized over the whole set of locations. Intuitively, it is because of the assumption that migration is costless. That is, the migration decision does not depend on the origin location, as moving between any location pair has zero cost.

Derivation of (1.6) from discrete time can be found in Appendix A. In words, agents derive a flow rate of utility of $A_i \varepsilon C_i^T(t)$ in location i. With a rate of ζ, they draw new tastes for locations, and decide to migrate after which they earn an expected life-time utility of $\bar{\mathcal{U}}^T(t)$. Thus the utility return from drawing new taste shocks is given by $\bar{\mathcal{U}}^T(t) - \mathcal{U}_i^T(\varepsilon, t)$. The last term stands for the time appreciation in the value function.

Finally, the equilibrium number of workers and inventors located in i are denoted by $L_i(t)$ and $R_i(t)$, respectively, satisfying that $\sum_i L_i(t) = \bar{L}$ and $\sum_i R_i(t) = \bar{R}$.

1.3.4 Production

There are two types of goods produced in the economy. Final good is used for consumption and as an input for intermediate good production. The price of final good is normalized to one. Intermediate goods are used as an input for the production of the final good. Workers are employed in final good sector in each location. All goods are shipped across locations without any costs. Thus, the place of production is not important apart from the fact that they create demand for labor.

Final goods. Production function of final good in location i is given by

$$Y_i(t) = \bar{\mathcal{A}} L_i(t)^\beta \left[\exp\left(\int_0^1 \log k_i(\nu, t) d\nu \right) \right]^{1-\beta} \tag{1.7}$$

All locations have access to the identical production technology (1.7). $\bar{\mathcal{A}}$ is a constant term to simplify algebra.[3] $L_i(t)$ denotes the amount of workers employed in location i, and $k_i(\nu, t)$ denotes the amount of intermediate good $\nu \in [0, 1]$ demanded by location i. The elasticity of substitution between factors of production is one, and β represents the share of labor in the production of final good.

3. Given the innovation step size parameter $\lambda > 1$ defined below, $\bar{\mathcal{A}} = \left(\frac{\lambda}{1-\beta}\right)^{1-\beta}$.

Final goods are produced in perfectly competitive markets. As trade is costless, the price of intermediate good ν is identical in all markets i. Denoting this price by $p(\nu, t)$, demand functions for production factors are given by

$$W_i^L(t)L_i(t) = \beta Y_i(t)$$
$$p(\nu, t)\, k_i(\nu, t) = (1-\beta)\, Y_i(t)$$

Intermediate goods and R&D. A unit measure of intermediate goods are differentiated, and each variety is denoted by $\nu \in [0, 1]$. These varieties are simply called products throughout the paper. Each product is produced by a single intermediate good firm (or simply firm) in equilibrium, however, a firm can produce multiple products. The intermediate good sector is identical to Klette and Kortum [2004]'s model of heterogeneous firms and innovation. In addition to the fact that firms are defined by the portfolio of products they produce, as in Klette and Kortum [2004], I make only one addition to their structure, i.e. firms establish their single R&D lab in a location, and they perform their R&D and innovation in this lab by hiring researchers. I assume that firms cannot change the location of the R&D lab throughout their life-time.

Intermediate goods are produced from final goods with a linear technology

$$k(\nu, t) = a(\nu, t)\, y(\nu, t)$$

where $a(\nu, t)$ is productivity, and $y(\nu, t)$ is the final good used in production. Unit elastic demand for intermediate goods implies that the firm that holds the most frontier technology for a good ν captures the whole market, and charges a markup so that the price charged is just equal to the marginal cost of the second most productive firm. As the price of the final good is normalized to one, the price of variety ν equals $p(\nu, t) = \lambda a(\nu, t)^{-1}$ where $\lambda > 1$ denotes the innovation step size as described below.

Firms invest in R&D to expand their product portfolio. Upon a successful innovation, the firm comes up with a new production technology for a random product $\nu' \in [0,1]$ drawn from a Uniform distribution on the unit interval, improving previous productivity $a\left(\nu', t\right)$ by a constant factor of $\lambda > 1$, and obtains the monopoly of the good with a higher productivity level of $\lambda a\left(\nu', t\right)$.[4] As the price of final good is one in all locations, the relative marginal cost of past incumbent to that of the current incumbent in any product line is always equal to λ. This is the reason as to why the maximum markup that can be charged equals to λ as described above. The flow of profits per unit of time is then equals to $\pi Y(t)$ for all products $\nu \in [0,1]$, where $\pi \equiv \left(1-\lambda^{-1}\right)(1-\beta)$ and $Y(t) \equiv \sum_n Y_n(t)$ is the the total output produced in the economy. As the rate of profit per product does not depend on the location of production, firms are indifferent in where to produce their products. The production locations of intermediate goods are indeterminate in the model because of the absence of trade costs, and the identical cost of production factor for all firms, i.e. the final good, no matter where they locate their production plant.

Let n denote the number of products owned by a firm, and i denote the location of the R&D lab of the firm. The R&D production technology of the firm is given by

$$Z_i\left(n,t\right) = \left[\alpha_i(t) R_i(n,t)\right]^{\frac{1}{\theta}} n^{1-\frac{1}{\theta}}, \qquad \theta > 1 \tag{1.8}$$

In equation (1.8), $Z_i(n,t)$ denotes the rate of innovation which is the Poisson arrival rate of a successful innovation, and R denotes the number of researchers employed in the R&D lab located in i. Researchers in location i benefit from the location research productivity $\alpha_i(t)$ as described in Section 1.3.1. Moreover, it is assumed that larger firms in terms of the number of products owned are more productive in R&D. The parameter θ governs the curvature of the innovation production function with respect to the number of inventors employed.

4. As all firms have a countable number of products in their portfolio, the probability of the event in which the firm invents on one of its products equals zero.

Firms are owned by all individuals in the economy. As they have linear preferences over time, the rate of interest equals to the time discount rate ρ. Thus, the HJB function for the value function of a firm with n products and located in i can be written as follows

$$\begin{aligned}\rho \mathcal{V}_i(n,t) - \dot{\mathcal{V}}_i(n,t) = & \underbrace{n\pi Y(t)}_{\text{Profit}} + \underbrace{nx(t)\left[\mathcal{V}_i(n-1,t) - \mathcal{V}_i(t)\right]}_{\text{Value loss from creative destruction}} \\ & + \max_R \Big\{ \underbrace{-\left(1-s_i\right) W_i^R(t)R}_{\text{Cost of R\&D}} + \underbrace{\left[\alpha_i(t)R\right]^{\frac{1}{\theta}} n^{1-\frac{1}{\theta}}}_{\text{Innovation rate } Z_i(n,t)} \underbrace{\left[\mathcal{V}_i(n+1,t) - \mathcal{V}_i(n,t)\right]}_{\text{Value gain from inn.}} \Big\}\end{aligned} \tag{1.9}$$

where $\mathcal{V}_i(n,t)$ denotes the sum of discounted future profits. The right hand side of (1.9) has multiple terms. The first one is the total rate of profits generated from the whole product portfolio. The second term is the loss in firm value due to creative destruction. $x(t)$ denotes equilibrium aggregate rate of creative destruction per product, which is defined as the total rate of innovation per unit of time from all locations in the economy. Firms take this as given when deciding how much to invest in R&D. As the firm has n products, the total rate of creative destruction that the firm faces equals to $nx(t)$. When other firms obtain a superior technology, the incumbent firm loses one of the products from its portfolio. If the firm had a single product, then the firm exits in such a case. The last term is the value stemming from the R&D investments. The first term in the maximization problem is the total cost of innovation which equals to the wage bill of researchers employed, after s_i portion of it is subsidized by the government. In particular, $W_i^R(t)$ denotes the equilibrium wage rate of inventors in location i, and $s_i \in [0,1]$ denotes the R&D subsidy rate in the location. The second term indicates the expected return from innovation which is the product of the rate of innovation and the value gain after firm adds an additional product to its portfolio.

Theorem 1.3.1. The solution to the HJB equation (1.9) has the form $\mathcal{V}_i\left(n,t\right) = n\nu_i(t)Y(t)$ where $v_i(t)$ denotes the normalized monopoly value of owning a product. Furthermore, per product value $v_i(t)$, per product innovation rate, defined as $z_i(t) \equiv \frac{Z_i(n,t)}{n}$, and per product inventor employment, defined as $r_i(t) \equiv \frac{R_i(n,t)}{n}$ are independent of the size of the firm's portfolio n and satisfy the following set of equations

$$z_i(t) = \left[\alpha_i(t) r_i(t)\right]^{\frac{1}{\theta}} \tag{1.10}$$

$$(1 - s_i)\, w_i^R(t) r_i(t) = \frac{1}{\theta}\left[\alpha_i(t) r_i(t)\right]^{\frac{1}{\theta}} v_i(t) \tag{1.11}$$

$$\dot{v}_i(t) = \left[\rho - g(t) + x(t) - \frac{\theta - 1}{\theta} z_i(t)\right] v_i(t) - \pi \tag{1.12}$$

where $w_i^R(t) \equiv \frac{W_i^R(t)}{Y(t)}$ denotes inventor wage in i normalized by aggregate output, and $g(t) \equiv \frac{\dot{Y}(t)}{Y(t)}$ is the growth rate of aggregate output.

Proof. See Appendix B. □

Theorem 1.3.1 states that the value of a product is independent of the size of the firm which is measured by n. This results stems from the Cobb-Douglass specification for the innovation production function (1.8).[5] Furthermore, per product line innovation rates and inventor employments are also independent of firm size. Equation (1.11) gives the demand for inventors by incumbent firms from a location. Ceteris paribus, high research productivity in a location causes higher demand for inventors for any given wage level. Therefore, $\alpha_i(t)$ can be regarded as a demand shifter for inventors across locations.

1.3.5 Entry

Location of R&D lab of incumbent firms is determined at the time of entry. There is a unit mass of potential entrants in the economy who are frictionlessly mobile across regions. Similar to incumbents, entrants employ researchers to generate a superior technology on a

5. A more detailed discussion on the implications of this R&D function can be found in Akcigit and Kerr [2018].

random product line $\nu \in [0,1]$. A potential entrant in location i who employs $\tilde{r}_i(t)$ inventors generates an innovation rate of $\tilde{z}_i(t)$ which is given by

$$\tilde{z}_i(t) = \frac{1}{f}\left[\alpha_i(t)\tilde{r}_i(t)\right]^{\frac{1}{\theta}} \tag{1.13}$$

The parameter f represents entry costs that are common to all locations, and the curvature parameter θ is same across entrants and incumbents. Importantly, inventors benefit from location specific research productivity $\alpha_i(t)$ whether they are employed by incumbent firms or entrants. Denoting the value of entry in location i by $\tilde{\mathcal{V}}_i(t)$, each potential entrant located in i solves the following entry problem

$$\tilde{\mathcal{V}}_i(t) \equiv \max_{\tilde{r}} \Big\{ \underbrace{-\left(1-s_i\right) W_i^R(t)\tilde{r}}_{\text{Cost of R\&D}} + \underbrace{\frac{1}{f}\left[\alpha_i(t)\tilde{r}\right]^{\frac{1}{\theta}}}_{\text{Inn. rate } \tilde{z}_i(t)} \underbrace{\mathcal{V}_i(1,t)}_{\text{Return}} \Big\} \tag{1.14}$$

The return from innovation for entrants is equal to the market value of an incumbent firm on the same location that starts with a single product.

Theorem 1.3.2. Let normalized value of being an entrant in location i be defined as $\tilde{v}_i(t) \equiv \frac{\tilde{\mathcal{V}}_i(t)}{Y(t)}$. Then,

$$\tilde{v}_i(t) = \frac{\theta-1}{\theta}\tilde{z}_i(t)v_i(t) \tag{1.15}$$

Furthermore, inventor employment of a potential entrant is proportional to that of incumbent firms in their location. As a result, per potential entrant innovation rate is also proportional to per product innovation rate of the location. That is,

$$\tilde{r}_i(t) = \frac{1}{F}r_i(t) \tag{1.16}$$

$$\tilde{z}_i(t) = \frac{1}{F}z_i(t) \tag{1.17}$$

where F is a composite parameter defined as $F = f^{\frac{\theta}{\theta-1}}$.

Proof. See Appendix C. $\square$

Theorem 1.3.2 can be proven easily by combining first order conditions to incumbent and entrant problems, (1.9) and (1.14). The implication of this theorem is that entrant choices are closely linked to incumbent firms in their location. The reason is that they have access to a similar R&D production function with incumbent firms, and inventors benefit the location specific R&D resources, $\alpha_i(t)$, both in incumbent and entrant firms in a location. This structure is particularly chosen so that entry part of the model simplifies considerably. The only mission of entrants in the model is to give rise to new firms that exit frequently due to creative destruction $x(t)$. However, entrants' location choice is not trivial. Indeed, they are indifferent in equilibrium between locations to perform R&D and enter to the market. Labeling this equilibrium condition as free entry condition across locations, it can be formally stated as[6]

$$\tilde{v}_i(t) = \tilde{v}_j(t), \qquad \forall i, j, t \tag{1.18}$$

The free entry condition (1.18) pins down the equilibrium mass of potential entrants across locations which are denoted by $\left\{\tilde{\psi}_i(t)\right\}_i$ such that $\sum_i \tilde{\psi}_i(t) = 1$ for all t. Similarly, the total measure of product lines owned by firms from location i in equilibrium is denoted by $\psi_i(t)$ such that $\sum_i \psi_i(t) = 1$. The variable $\psi_i(t)$ has an endogenous evolution over time as a result of firm innovation choices and entry rates in the location. Intuitively, it increases in the number of potential entrants in i, $\tilde{\psi}_i(t)$ as more entry means higher survival rate of location i firms compared to other regions. Formal derivations are delegated to Appendix D.

6. Normally, this condition should be stated at the non-normalized levels of entry values, i.e. $\tilde{\mathcal{V}}_i(t) = \tilde{\mathcal{V}}_j(t)$ $\forall i, j, t$.

1.3.6 Knowledge diffusion across locations

As described in Section 1.3.1, research productivity in locations depends on endogenous flow of past ideas within the country. I assume that each innovation embeds a measure of ideas normalized to one. After the invention of these ideas in an origin location j, they diffuse to the rest of the economy in order to lay the foundation for the new ideas to be invented, possibly combined with other ideas that have diffused from somewhere else. However, the diffusion process is not homogeneous and perfect across location pairs. I assume that ideas diffuse between locations with a random time lag. Specifically, let $\omega_{ij} > 0$ be called the rate of diffusion from j to i. Then, the time lag for which an idea originated in location j diffuses to location i is a random variable distributed as Exponential $\left(\omega_{ij}\right)$. The parameter of this distribution, ω_{ij}, varies across origin-destination pairs, and it is possible that $\omega_{ij} \neq \omega_{ji}$. It follows from this structure that the mean time lag of idea diffusion from j to i is equal to $1/\omega_{ij}$. As $\omega_{ij} \to 0$, ideas never diffuse from j to i in finite time. On the contrary, as $\omega_{ij} \to \infty$, the diffusion becomes instantaneous. The $N \times N$ matrix $\Omega = \left[\omega_{ij}\right]$ holds the diffusion rate parameters, and it is called the knowledge network throughout the paper.

In line with the empirical evidence in Section 1.2.2, it is also assumed that ideas get obsolete at an exogenous rate of $\delta > 0$ over time.[7] As ideas get older, they are more likely to be replaced by new and better ideas over time.

In order to derive the evolution of $K_i(t)$, where i is called the destination location, it is required to define a variable which represents the number of ideas that are invented in j, but have not yet diffused to i by time t. This variable is denoted by $K'_{ij}(t)$. Then, the law of motion of $K_i(t)$ can be derived as

$$\dot{K}_i(t) = \sum_{j=1}^{N} \omega_{ij} K'_{ij}(t) - \delta K_i(t) \tag{1.19}$$

7. Obsolescence of old ideas can be endogenized by the creative destruction process of frontier technologies.

The first term on the right hand side is a summation across all locations. The flow of ideas to location i from j is equal to the rate of diffusion times the stock of ideas available for diffusion. On the other hand, ideas get obsolete over time with a rate of δ, which is captured by the second term on the right hand side of (1.19).

How does $K'_{ij}(t)$ evolves over time? Let $x_j(t)$ be the total rate of innovation in origin location j. Then, it is equal to $x_j(t) = \psi_j(t)z_j(t) + \tilde{\psi}_j(t)\tilde{z}_j(t)$. The inflow to the stock of not-yet-diffused ideas to a particular location i equals the number of ideas embedded in an invention, which is normalized to one, times the rate of innovation, $x_j(t)$. On the other hand, the outflow of ideas from this stock is due to either diffusion to i or obsolescence. Hence, we can show that $\dot{K}'_{ij}(t)$ equals to

$$\dot{K}'_{ij}(t) = x_j(t) - \left(\omega_{ij} + \delta\right) K'_{ij}(t) \tag{1.20}$$

Equation (1.19) suggests that the rate at which $K_i(t)$ grows over time increases with ω_{ij}, and the size of the stock of ideas waiting to be diffused, $K'_{ij}(t)$. This stock, on the other hand, is positively correlated with the rate of innovation j, i.e. $x_j(t)$. Therefore, connectedness represented by ω_{ij} is not the only determinant of the size of past ideas available for use in i. Locations that are particularly connected to innovation hubs, i.e. locations with high $x_j(t)$, benefit from knowledge spillovers relatively more.

1.3.7 Market clearing conditions

Workers. Supply of workers in a location, determined from their migration decision (1.4), is equal to the labor demand from final good producers in the location. Clearing intermediate good markets along with local worker markets in each location gives rise to a very simple solution for the wage rate of workers. Delegating the derivations to Appendix E, we can show that worker wage rate is common across locations, and is given by

$$W_i^L(t) = W^L(t) = \frac{\beta}{\bar{L}} Y(t), \qquad \forall i = 1, \ldots, N \tag{1.21}$$

In the absence of trade costs, the marginal productivity of workers across locations are equal, hence they earn equal wages, which is proportional to aggregate output at all times. Thus, when workers move across locations, they only value relative amenities. This implication of the model allows me to control for amenity differences across locations by matching observed worker allocations in space. Therefore, the remaining variation in inventor-to-worker ratio across locations informs relative inventor wages, which is heterogeneous across locations.

Total output of the economy equals to $Y(t) = \mathcal{A}(t)^{\frac{1-\beta}{\beta}} \bar{L}$, where $\mathcal{A}(t)$ is the aggregate productivity index defined by

$$\mathcal{A}(t) \equiv \exp\left[\int_0^1 \log a(\nu, t) d\nu\right] \tag{1.22}$$

which is a unit elastic aggregation across productivity of all intermediate goods. The source of growth stems from innovations at the intermediate goods level. Furthermore, as shown in Appendix F, the growth rate of aggregate productivity equals to $\frac{\dot{\mathcal{A}}(t)}{\mathcal{A}(t)} = \log(\lambda)\, x(t)$. Thus, the growth rate of output is given by

$$g(t) = \frac{1-\beta}{\beta} \log(\lambda)\, x(t) \tag{1.23}$$

Inventors. Total supply of inventors in a location, $R_i(t)$, is determined by inventor migration choice given by (1.4). The demand, on the other hand, is equal to total inventor employment in a location, which is the sum of incumbent's and entrant's demand for inventors. Hence, market clearing condition for inventors in location i can be stated as

$$R_i(t) = \psi_i(t) r_i(t) + \tilde{\psi}_i(t) \tilde{r}_i(t) \tag{1.24}$$

Government budget constraint. Government finances location specific R&D subsidies from the taxation of individuals' labor income. It is assumed that it holds period-by-period

$$\sum_{i=1}^{N} s_i W_i^R(t) R_i(t) = \sum_{i=1}^{N} \tau(t) W_i^L(t) L_i(t) + \sum_{i=1}^{N} \tau(t) W_i^R(t) R_i(t)$$
$$\implies \tau(t) = \frac{\sum_{i=1}^{N} s_i W_i^R(t) R_i(t)}{\sum_{i=1}^{N} W_i^L(t) L_i(t) + W_i^R(t) R_i(t)} \tag{1.25}$$

Total profits and their allocation across agents are derived in Appendix G.

1.3.8 Equilibrium and predictions of the model

We can now proceed with the equilibrium properties and the predictions of the model on equilibrium wage rate of inventors and inventor allocation across locations. The particular equilibrium that is considered in the paper is the balanced growth path (BGP) equilibrium in which the growth rate of the economy $g(t)$ is constant over time. Moreover, in this equilibrium, the growth rate of inventor wages in all locations are equal to the growth rate of output. Thus, the model variables stay stationary in this equilibrium after normalizing them with the aggregate output $Y(t)$. In what follows, I will conjecture that the model admits a BGP equilibrium, derive its predictions, and then finally show that the initial conjecture holds.

Knowledge network and research productivity. In BGP, location innovation rates x_j are time invariant. Under this conjecture, the system of differential equations given by (1.19) and (1.20) have a stationary solution given by $K_i = \frac{1}{\delta} \sum_{j=1}^{N} \frac{\omega_{ij}}{\omega_{ij}+\delta} x_j$. Thus location research productivity $\alpha_i = \bar{\alpha}_i^{1-\varphi} K_i^{\varphi}$ is also constant over time. Moreover, the rate of creative destruction which is the total innovation rate in the economy is a constant and equals to $x = \sum_i x_i$.

Innovation rates. Replacing the aggregate creative destruction rate x into (1.23) implies that the growth rate of the economy equals to $g = \frac{1-\beta}{\beta} \log(\lambda) x$. The second conjecture of the BGP equilibrium is that the normalized inventor wage w_i^R and inventors per product r_i are constant over time. Under these conjectures, we can show that $\dot{v}_i(t) = 0$, and stationary values z_i, r_i and v_i satisfy the system of equations given by (1.10), (1.11) and (1.12). Importantly, we have

$$v_i = \frac{\pi}{\rho - g + x - \frac{\theta-1}{\theta} z_i} \tag{1.26}$$

Equation (1.15) combined with free entry condition (1.18) implies that per product rate of innovation in locations are equal to a common rate, $z_i = z_j = z$ for all i, j. Using the relationship between incumbent and entrant innovations given by equation (1.17), we also have that entrants in all locations choose the same rate of innovation $\tilde{z}_i = \tilde{z} = z/F$.

Theorem 1.3.3. Let ψ_i denote the total measure of product lines owned by incumbent firms that are located in i, and let $\tilde{\psi}_i$ denote the measure of potential entrants located in i. Then, in BGP equilibrium,

$$\psi_i = \tilde{\psi}_i = \frac{\alpha_i R_i}{\sum_{j=1}^{N} \alpha_j R_j} \tag{1.27}$$

Proof. See Appendix H. □

Although firms and entrants choose equal rates of innovations in any location, the difference between regions in terms of innovative activity stems from the extensive margin. That is, in equilibrium, firms located in more research productive regions obtain a higher share of market ownership which is measured by the mass of product lines owned by local firms. The intuition is as follows. In equilibrium, more entrants prefer high research productive locations. Therefore, the entry rate in those locations are higher. Since all firms in the economy face the same exit probability which is implied by the aggregate creative destruction rate x, a startup cohort from more productive locations are more successful in surviving in the market because of their large population due to high entry rate. Thus, in equilibrium, firms from research productive locations survive better and capture a larger fraction of product markets in overall economy. Although I do not test the spatial firm dynamics predictions of the model in this paper, the difference across locations stemming from heterogeneous firm dynamics allows me to explain the fundamental source of high demand for inventors in certain locations. In other words, in the model, the reason for high inventor demand in certain locations is not directly due to the presence of high volume of innovative firms there. Instead, there is another factor, endogenous research productivity of locations, which gives rise to both phenomena simultaneously, i.e. high inventor demand and high number of innovative firms.

Inventor wage across locations. Another important prediction of the model is that inventor wages are proportional to research productivity of locations. Next theorem shows this result

Theorem 1.3.4. Let w_i^R denote the inventor wage rate in equilibrium normalized by aggregate output. Then

$$w_i^R = \frac{1}{\theta} z^{1-\theta} v \frac{\alpha_i}{1 - s_i} \tag{1.28}$$

where z is per product line innovation rate common to all locations, and $v = \frac{\pi}{\rho - g + x - \frac{\theta - 1}{\theta} z}$ following from (1.26), and $\alpha_i = \bar{\alpha}_i^{1-\varphi} K_i^{\varphi}$ is the research productivity of location i.

Proof. See Appendix I. □

Equilibrium inventor wage in a location increases with research productivity of the location and the subsidies provided for R&D activities. This prediction of the model allows me to pin down relative research productivities of locations by exactly matching inventor allocation across US states. Next section describes the migration behavior of agents in BGP, and shows the resulted allocation of workers and inventors across space.

Migration and inventor allocation in space. In order to simplify migration problem of agents given by (1.4), we first need to solve the value function $\mathcal{U}_i^T(\varepsilon, t)$ in BGP.

Theorem 1.3.5. Agent value function $\mathcal{U}_i^T(\varepsilon, t)$ as defined in Section 1.3.3 has an analytical solution in BGP equilibrium as follows

$$\mathcal{U}_i^T(\varepsilon, t) = \frac{A_i \varepsilon C_i^T(t)}{\rho + \zeta - g} + \frac{\zeta}{(\rho + \zeta - g)(\rho - g)} \Gamma\left(1 - \frac{1}{\xi}\right) \left[\sum_{j=1}^{N} \left(A_j C_j^T(t)\right)^{\xi}\right]^{\frac{1}{\xi}} \tag{1.29}$$

where $\Gamma(\cdot)$ is Gamma function, and $C_i^T(t) = (1 + d - \tau) w_i^T Y(t)$ is consumption of type-T in i that is proportional to aggregate output. Thus, $\mathcal{U}_i^T(\varepsilon, t)$ is also proportional to $Y(t)$.[8]

Proof. See Appendix J. □

Having equipped with agent values, the migration choice (1.4) simplifies considerably in BGP as stated by the next theorem.

Theorem 1.3.6. Let $\left(i^T\right)^\star$ be the location choice of agents of type-T in BGP, conditional on a set of location tastes given by vector $\mathbf{e}$. Then,

$$\left(i^T\right)^\star = \arg\max_j \left\{A_j e_j w_j^T\right\} \tag{1.30}$$

where e_j is the j$^{\text{th}}$ component of $\mathbf{e}$. Furthermore, let γ_i^T be the fraction of type-T population located in i. Given worker and inventor wages in (1.21) and (1.28), the migration choice

8. The analytical expressions of d and τ in BGP equilibrium are given in proof.

(1.30) implies

$$\gamma_i^L = \frac{A_i^{\xi}}{\sum_{j=1}^{N} A_j^{\xi}} \tag{1.31}$$

$$\gamma_i^R = \frac{\gamma_i^L \left(\frac{\alpha_i}{1-s_i}\right)^{\xi}}{\sum_{j=1}^{N} \gamma_j^L \left(\frac{\alpha_j}{1-s_j}\right)^{\xi}} \tag{1.32}$$

Thus, number of workers and inventors in locations can be found as $L_i = \gamma_i^L \bar{L}$ and $R_i = \gamma_i^R \bar{R}$.

Proof. See Appendix K. □

Theorem 1.3.6 forms the basis for the identification of location specific research productivities. Heterogeneity in amenities across locations, which is an important ingredient in inventor supply to local labor markets, is controlled for by observed worker allocation in space. Simple structure of the model aggregates possibly many different characteristics of locations under a single residual, the amenity A_i. The assumption needed is that any such characteristics affect both types of individuals identically when they decide where to relocate. Further implication of (1.32) is that inventor-to-worker ratio in a location increases with its research productivity. That is,

$$\frac{\alpha_i}{\alpha_j} = \frac{1-s_i}{1-s_j} \left[\frac{\left(\gamma_i^R / \gamma_i^L\right)}{\left(\gamma_j^R / \gamma_j^L\right)} \right]^{\frac{1}{\xi}} \tag{1.33}$$

Equation (1.33) identifies relative research productivity of locations given taste dispersion parameter ξ, as the right hand side of the equation is observable in the data.[9] Given worker and inventor allocations from the data, we can solve for relative research productivity. As

9. As discussed in Section 1.4, observed inventor allocation in patent data is not exactly equal to true inventor allocation, as the data only consists of the inventors that applied for a patent in a given period of time. As will be shown in the same Section, the model structure allows us to make a connection from the number of "successful" inventors who applied for a patent to true number of inventors including "unsuccessful" ones.

$\alpha_i = \bar{\alpha}_i^{1-\varphi} K_i^{\varphi}$ holds true, we can further decompose α_i in fundamental research productivity of location, $\bar{\alpha}_i$, and network effects captured by K_i, as explained below.

Verifying initial conjectures. The predictions derived up to this point depend on the initial conjectures that x_i, w_i^R and r_i are constant over time. Theorem 1.3.4 proves that w_i^R is indeed constant. From equation (1.8), we can show that $r_i = z^{\theta}/\alpha_i$, which does not vary over time. Following theorem proves that x_i is also a constant in BGP equilibrium.

Theorem 1.3.7. In BGP, the total rate of innovation in a location x_i and per product innovation rate z can be derived as follows

$$z = \left[\frac{F}{1+F} \sum_{i=1}^{N} \alpha_i R_i \right]^{\frac{1}{\theta}} \tag{1.34}$$

$$x_i = z^{1-\theta} \alpha_i R_i \tag{1.35}$$

Further replacing x_i in $x = \sum_i x_i$ implies that the aggregate rate of creative destruction equals to

$$x = \frac{1+F}{F} z \tag{1.36}$$

Thus, aggregate growth rate of the economy is finally

$$g = \frac{1-\beta}{\beta} \log(\lambda) \frac{1+F}{F} z \tag{1.37}$$

which is proportional to z.

Proof. See Appendix L. □

Theorem 1.3.7 verifies the initial conjecture that x_i are constant over time. Moreover, equation (1.35) and definition $\alpha_i = \bar{\alpha}_i^{1-\varphi} K_i^{\varphi}$ result in a nonlinear system of equations in

$\{K_i\}_i$ such that

$$K_i = \frac{1}{\delta z^{\theta-1}} \sum_{j=1}^{N} \frac{\omega_{ij}}{\omega_{ij}+\delta} \bar{\alpha}_j^{1-\varphi} K_j^{\varphi} R_j \tag{1.38}$$

Equation (1.38) is endogenous in the sense that knowledge spillovers across locations depend on inventor allocation through their effect on innovation intensity in locations. Ceteris paribus, regions that are more connected to the locations with large inventor populations benefit more from knowledge spillovers. The reason is the direct effect of inventor population on idea creation in origin locations. More inventors create more ideas per unit of time, and these ideas spill to other connected locations faster. Furthermore, fundamental research productivity $\bar{\alpha}_j$ and inventors R_j reinforce this effect, as inventors are more likely to migrate to locations with high $\bar{\alpha}_j$.

1.3.9 Social welfare function and planner's problem

In this section, social welfare function is derived based on agent value functions found in Theorem 1.3.5. It is assumed that the social planner cares about the ex-ante expected value of agents, $\bar{\mathcal{U}}^T(t)$, before they draw idiosyncratic taste shocks and migrate to the location that provide the highest value for themselves, which is given by definition (1.5). The function $\bar{\mathcal{U}}^T(t)$ represents the social welfare of type-T agents because the planner internalizes agents' migration decisions based on their idiosyncratic location preferences, and she knows that they would migrate to the highest value locations. It is a good choice in comparing long run equilibria under different counterfactuals, as it abstracts away from transition periods during which agents relocate across locations between regions.

Utilizing the analytical expression for the agent values given by equation (1.29) and the fact that idiosyncratic location tastes ε_i are drawn from Frechet distribution, we can derive

$\bar{\mathcal{U}}^T(t)$ as follows[10]

$$\bar{\mathcal{U}}^T(t) = \Gamma\left(1 - \frac{1}{\xi}\right)\frac{1}{\rho - g}\left[\sum_{i=1}^{N}\left(A_i C_i^T(t)\right)^{\xi}\right]^{\frac{1}{\xi}} \tag{1.39}$$

The derivation of this expression depends on the convenient properties of the Frechet distribution, and can be found in Appendix J. It should be noted that this expression is independent of migration frequency parameter ζ, since the function $\bar{\mathcal{U}}^T(t)$ represents the value of agents independent of their initial locations. Secondly, the welfare of agents increase with the aggregate growth rate of the economy, as higher growth implies higher consumption in the future. Finally, $\bar{\mathcal{U}}^T(t)$ can be considered as a weighted average of location specific consumption rates, weights being the amenities in locations A_i. The planner cares about an aggregate consumption across all locations, however, consumption in high amenity locations are valued relatively more.

The final social welfare function is defined as a weighted average of worker and inventor welfares, where weights are chosen by the planner. It is defined as follows

$$\begin{aligned}\mathcal{W}(t) &\equiv \phi^L \bar{L} \times \bar{\mathcal{U}}^L(t) + \phi^R \bar{R} \times \bar{\mathcal{U}}^R(t) \\ &= \Gamma\left(1 - \frac{1}{\xi}\right)\frac{1}{\rho - g}\left\{\phi^L \bar{L}\left[\sum_{i=1}^{N}\left(A_i C_i^L(t)\right)^{\xi}\right]^{\frac{1}{\xi}} + \phi^R \bar{R}\left[\sum_{i=1}^{N}\left(A_i C_i^R(t)\right)^{\xi}\right]^{\frac{1}{\xi}}\right\}\end{aligned} \tag{1.40}$$

The welfare weights for different agent types are given by ϕ^L and ϕ^R such that $\phi^L + \phi^R = 1$. In the rest of the paper, these weights are taken equal to each other, i.e. $\phi^L = \phi^R = 0.5$.

The planner maximizes the social welfare function (1.40) by choosing location specific

10. This derivation implicitly assumes that the growth rate of the economy g always stays lower than the time discount rate of ρ. Otherwise, agent value function explodes to infinity as the future consumption growth rate is higher than the discount rate.

R&D subsidy rates $s_i \in [0,1]$ subject to equilibrium condition in BGP. That is, the planner solves her problem in a constrained environment with a single policy tool available to her, place-based R&D subsidy rates. Taxation of R&D expenditures are not considered as a policy tool. In order to simplify the analysis, the tax rate chosen to finance the cost of the policy is set uniformly across all the locations. That is, while subsidy rates are location specific, labor income tax rate τ is uniform across locations. As the knowledge spillovers between locations has a nonlinear form given by (1.38), I solve planner's problem numerically.

Formally, the planner's problem can be stated as follows

$$\max_{s_i \in [0,1]} \Gamma\left(1-\frac{1}{\xi}\right)\frac{1}{\rho-g}\left\{\phi^L \bar{L}\left[\sum_{i=1}^{N}\left(A_i C_i^L(t)\right)^{\xi}\right]^{\frac{1}{\xi}} + \phi^R \bar{R}\left[\sum_{i=1}^{N}\left(A_i C_i^R(t)\right)^{\xi}\right]^{\frac{1}{\xi}}\right\} \tag{1.41}$$

s.t. BGP equilibrium conditions

1.4 Quantification

The model parameters are quantified with a combination of three steps. First of all, several aggregate parameters that are common to all locations are externally calibrated. Secondly, knowledge network represented by the matrix Ω is estimated from patent citation flows and citation lags between US states, which are the geographic unit of the analysis. Finally, I use the model to recover the remaining location specific parameters—fundamental research productivity $\bar{\alpha}_i$, and amenity A_i—from data on worker and inventor allocations in the US. Another aggregate parameter, entry cost f, is recovered from exactly matching model implied entry rate and the data counterpart. The method, which I call model inversion, infers location specific parameters that deliver worker and inventor location choices across US states as equilibrium outcomes.

The main intuition behind the estimation procedure outlined above is based on two important predictions of the model, given by equations (1.33) and (1.38). The first equation states that equilibrium level of endogenous research productivity of locations, a combination of exogenous factors and knowledge spillovers from other locations, can be inferred from relative ratio of inventor-to-worker fractions across locations. This result depends on the main assumption of the model—both inventors and workers value location amenities identically. After controlling for observed distribution of workers across US states, the remaining variation in inventor allocation identifies other factors that only affect inventors in their migration decisions, i.e. inventor wages. It should be noted that the implication of the simple structure of the model that workers earn the same wage in each location does not alter this line of reasoning. Even if worker wages were heterogeneous in a more complex model with location and worker specific productivity differences across locations, such a structural model would have allowed us to control for them via the corresponding migration decisions. The important point in this type of analysis is to have a structural model that would explain heterogeneous effects of locations on the earnings of different types of agents in the economy—workers and inventors. The model is intentionally kept simple for worker earnings characteristics so that the main intuition for the identification of heterogeneous research productivity of locations is more explicit.

The second equation (1.38) describes knowledge flows between locations. The main ingredient of this equation is diffusion rate parameters ω_{ij} which are specific to each state-pair. These parameters are inferred from patent citations and citation lags between state-pairs. The main assumption that justifies this exercise is a strong one, which is patent citations, although not perfect, reflect intertemporal knowledge spillovers in the innovation process. Citing inventors cite previous inventions from which they learn and inspire, and on which build. Thus, availability of this knowledge, K_i, increases their research productivity, as in the model.

Externally calibrated parameters. Table 1.1 gives the list of externally calibrated parameters and the corresponding values. Except place-based R&D subsidies, which are taken

Table 1.1: Externally calibrated parameters

Parameter	**Description**	**Value**	**Source**
ρ	Time discount rate	0.05	Matching 5% annual real rate
β	Labor share in production	0.6	Labor share
λ	Innovation step size	1.15	General literature
θ	Curvature of innovation function	2	General literature
δ	Idea obsolescence rate	0.075	Caballero and Jaffe [1993]
ξ	Location taste dispersion	2	Desmet et al. [2018]
φ	Share of past knowledge in research prod.	0.5	Externally set
$\bar{L}$	Total mass of workers	1	Normalization
s_i	R&D subsidy rate	0	Externally set

to be zero, none of externally calibrated parameters vary across locations. One of the important parameters in this list is location taste dispersion (ξ) which directly affects endogenous sorting of agents into locations. In particular, although it does not alter the ranking of locations for estimated parameters, dispersion parameter shapes the concentration of agents in equilibrium. Lower values for ξ means that agents have more dispersed preferences for locations, hence in equilibrium less concentration arises. Another important parameter is φ that governs the importance of intertemporal knowledge spillovers in the innovation process relative to other location-specific exogenous factors. Higher φ corresponds to a higher share of past knowledge in the creation of future inventions. In what follows, this parameter is taken to be half, i.e. $\varphi = 0.5$.

The time frequency of the model is taken to be a year. As agents have linear preferences over time, the equilibrium interest rate equals to ρ which is taken to be 5% (annual), which is common in endogenous growth literature. Innovation step size $\lambda = 1.15$ lies in the range of several estimates in the literature. This parameter mainly affects the growth rate of the economy suggested by equation (1.23). β is taken to be 0.6 which is in line with an average labor share of 60% for the period studied. Innovation curvature parameter θ affects the marginal cost of innovation through the curvature of R&D production function with respect to researchers employed. In other types of growth models in which individuals are sorted between production and R&D, the curvature parameter has a direct effect on the aggregate growth rate of the economy through allocation of total labor into research activity. However, in this model, the total supply of inventors is assumed fixed, therefore, such implications are absent. Finally, exogenous rate of idea obsolescence δ is taken from Caballero and Jaffe [1993] where they estimate a similar citation equation given by (1.42) (will be explained below) in order to estimate the extent of intertemporal spillovers between time periods. A value of $\delta = 0.075$ implies an idea obsolescence rate of 7.5% in a year. The sole magnitude of this variable has a direct effect on the growth rate of the economy as higher obsolescence rate reduces the effectiveness of past ideas on research productivity and growth.

It should be noted that the parameters that affect the aggregate growth rate of the economy only alters the overall level of location-specific research productivity estimates. For instance, a higher δ implies lower growth all else equal. In order to match the constant growth rate of 1.37%, model inversion results in higher level of location specific research productivities without altering relative research productivities across locations. A more detailed discussion on the identification of relative research productivities can be found below.

1.4.1 Estimation of knowledge network

In this section, I derive an equation of citation probabilities across locations exploiting the idea diffusion structure of the model. This equation is labelled as citation equation in the rest of the paper, and estimated from patent citations data. We start with a thought experiment by asking what is the probability of a patent issued in j at time s being cited by patents issued in i at a later date $t \geq s$? First of all, ideas become obsolete over time with a rate of δ. Assuming that every patent is embedded with an idea intensity of one (normalization), the number of useful ideas remaining in the patent by time t is given by $e^{-\delta(t-s)}$. This is the average fraction of ideas that remain from time s to time t under the assumption that ideas are subject to independent obsolescence shocks with a rate of δ per unit of time.[11] Secondly, I assume that inventors in location i at time t cite the patent if and only if they observe the idea in their location by time t. Equivalently, the necessary and sufficient condition for citation is the diffusion of the idea from j to i between time points s and $t \geq s$. Under the assumption of Exponential distribution of diffusion lags, the probability that an idea diffuses from j to i by time t is given by $1 - e^{-\omega_{ij}(t-s)}$. As diffusion and obsolescence are independent events, the probability of citation is given by the product of two probabilities, i.e. $e^{-\delta(t-s)}\left[1 - e^{-\omega_{ij}(t-s)}\right]$.

Citation probability is affected by a number of factors. All else equal, the time lag has two opposing effects on citation probabilities. Citation probability is negatively correlated with time lag $t - s$ because of idea obsolescence channel. As ideas age older, the probability that idea stays useful by time t declines (the first term). On the other hand, citation probability increases by the time lag, as it is more likely for ideas to be diffused to other locations as more time passes since their invention (second term). Other factors are due to parameters δ and ω_{ij}. All else equal, higher rate of obsolescence decreases citation probabilities between all location pairs uniformly. Finally, as idea diffusion rate ω_{ij} increases between locations, then it becomes more likely that patents from the destination location i cites past patents that originated from the origin location j in a fixed time interval of length $t - s$.

Under these assumption we can derive the maximum time lag at which patent citation probability is maximized. Taking first order condition of $\max_\tau e^{-\delta\tau}\left[1 - e^{-\omega_{ij}\tau}\right]$ with respect to τ yields a location pair specific time lag $\tau^\star_{ij}$ at which citation probability from i to j is maximized as follows, $\tau^\star_{ij} = \frac{1}{\omega_{ij}} \ln\left(\frac{\omega_{ij}+\delta}{\delta}\right)$. $\tau^\star_{ij}$ decreases with ω_{ij} which implies that the peak citation probability is reached earlier as the rate of diffusion between locations rises. This observation forms the basis for the identification of diffusion rates from patent citation

11. In the estimation procedure, δ is chosen as an annual rate of obsolescence of ideas

lags.

In order to derive an estimating equation of diffusion rates, I augment citation probability with location-time fixed effects separately for citing and cited locations. Let $\Gamma_{it}^{\text{citing}}$ denote the fixed effect for citing location i at time t which represents the technology composition of citing patents issued at time t in i. Similarly, $\Gamma_{js}^{\text{cited}}$ denotes the fixed effect for the technology composition of cited locations. These fixed effects aim to control for citing-cited technology composition of patent portfolios and its effect on the level of citation probabilities. The identification of ω_{ij} comes mainly from the citation lags. With the inclusion of fixed effects, the estimating equation becomes

$$\frac{C_{ij}^{ts}}{P_{it}P_{js}} = \Gamma_{it}^{\text{citing}} \times \Gamma_{js}^{\text{cited}} \times e^{-\delta(t-s)} \left[1 - e^{-\omega_{ij}(t-s)}\right] \tag{1.42}$$

The left hand side of equation (1.42) denotes the estimated patent citation probability which is defined as the observed number of citations from i's patents in time t to j's patents issued in time s (C_{ij}^{ts}) divided by the total number of all possible combinations between these two groups of patents, i.e. the product of the number of patents that are issued at time t in i (P_{it}) and the number of patents that were issued at time s in j (P_{js}). This equation is just equal to Caballero and Jaffe [1993] and Cai et al. [2022]'s citation equations, the only difference being it is modified in terms of citing-cited locations. In Caballero and Jaffe [1993], citing and cited fixed effects are included for time periods t and s in order to capture different number of ideas generated in these time periods and compositional differences as discussed above. They mainly focus on intertemporal spillovers between time periods by estimating a single diffusion rate parameter. In this paper, I estimate location-pair diffusion parameters by adding a spatial aspect to their citation equation. I estimate equation (1.42) by nonlinear least squares with an iterative minimization procedure by fixing the obsolescence parameter $\delta = 0.075$ as estimated by Caballero and Jaffe [1993].

1.4.2 Model inversion

Table 1.2: Targeted moments and identified parameters

Target	**Target Notation**	**Identified Parameter**	**Parameter Notation**
1. Allocation of workers across locations	$\gamma_i^L,\ i = 1, \ldots, N$	(Relative) amenity	$A_i,\ i = 2, \ldots, N$ with $A_1 = 1$
2. Allocation of patenting inventors across locations	$\gamma_i^{R\star} = \gamma_i^R,\ i = 1, \ldots, N$	(Relative) exogenous research prod.	$\bar{\alpha}_i,$
3. Total number of patenting inventors	$\bar{R}^\star = \frac{F^2+1}{F^2+F} z\bar{R}$	Total number of inventors	$\bar{R}$
4. Aggregate entry rate	$\tilde{z}$	Entry cost	f
5. Aggregate growth rate	$g = \frac{1-\beta}{\beta}\log(\lambda)x$	Level of exogenous research prod.	$\bar{\alpha}_1$

Equipped with the estimates of state-pair diffusion rates ω_{ij} from patent citation lags, we can finalize the quantification of the model with an inversion process by which location specific parameters, $\bar{\alpha}_i$ and A_i, are recovered. Table 1.2 shows the target moments from the data used, and the corresponding identified parameters. First of all, worker allocation across US states is targeted in order to pin down amenity distribution across locations. The mapping between the two is given by equation (1.31). Denoting the number of locations by N, we have $N-1$ many moments to match, as the sum of fractions of workers across states adds up to one. Therefore, we can recover location amenities only up to a scale. After normalizing the amenity in the first location to be one, i.e. $A_i = 1$, the equation (1.31) implies

$$A_i = \left(\frac{\gamma_i^L}{\gamma_1^L}\right)^{\frac{1}{\xi}}$$

Higher relative worker share in a location suggests higher level of amenities in the location, as worker wages are equalized across locations. Thus, the only heterogeneity remaining in worker migration decisions stems from location specific characteristics A_i.

Secondly, observed inventor allocation across US states is utilized to estimate fundamental research productivities of locations, $\{\bar{\alpha}_i\}_{i=1}^N$, only up to a scale. As discussed above, observed number of inventors in the patent data cannot be directly mapped to the number of inventors in the model, as not all inventors apply for a patent in a given year. To back out the true inventor allocation from the data, I utilize the model's predictions on innovation probabilities.

In the model, the rate of probability of patenting (or innovation) per product line is given by zdt, where dt is the length of the time interval considered, which is taken as one year. As per product inventor employment in location i is given by r_i, the number of inventors that come up with a new invention per product line in a time interval of dt is equal to $zdt \times r_i$. Similarly, $\tilde{z}dt \times \tilde{r}_i$ many inventors who are employed by entrants are successful in patenting. Denoting the total number of successful inventors in i as $R_i^\star$, and taking $dt = 1$, we have

$$R_i^\star = \psi_i z r_i + \tilde{\psi}_i \tilde{z} dt \tilde{r}_i = \left(1 + \frac{1}{F^2}\right) z \psi_i r_i$$

From inventor market clearing (1.24), we also have

$$R_i = \psi_i r_i + \tilde{\psi}_i \tilde{r}_i = \frac{F+1}{F} \psi_i r_i \implies \psi_i r_i = \frac{F}{F+1} R_i$$

Substituting this expression into the first one yields a relationship between $R_i^\star$ and R_i such that

$$R_i = \frac{F^2+F}{F^2+1} \frac{1}{z} R_i^\star \implies \bar{R} = \frac{F^2+F}{F^2+1} \frac{1}{z} \bar{R}^\star$$

where $\bar{R}^\star$ is the number of successful inventors nationwide, and the second expression follows from the summation of the first across all locations. Therefore, for a given estimate of $F = f^{\theta/(\theta-1)}$ and model implied z, we can map observed $\bar{R}^\star$ to the unknown $\bar{R}$. Lastly, the fraction of all inventors located in i in the model, γ_i^R, equals to i's share of successful inventors, $\gamma_i^{R\star} \equiv R_i^\star / \bar{R}^\star$. This can be seen by dividing both equations to each other which gives rise to $\gamma_i^R = \gamma_i^{R\star}$. Note that $\gamma_i^{R\star}$ and $\bar{R}^\star$ are observed from the patent data.

Fundamental research productivities $\bar{\alpha}_i$ are recovered in two steps. Firstly, equation (1.33) is used to back out endogenous relative productivities α_i, after replacing γ_i^R with $\gamma_i^{R\star}$. The intuition is explained as before, i.e. after controlling for amenity differences by the observed worker allocations, we can then estimate other factors that alter inventor earnings across locations. These factors are captured by location level resources for R&D and innovation. Secondly, we can use the definition $\alpha_i = \bar{\alpha}_i^{1-\varphi} K_i^{\varphi}$ and the endogenous formation of K_i across locations using model implied location innovation rates. In particular, an iterative procedure is employed using equation (1.38) given estimates of α_i to recover exogenous research productivities across locations $\bar{\alpha}_i$ up to a scale.

The aggregate entry rate in the model is equal to $\tilde{z}$, as all potential entrants choose the same innovation rate no matter where they are located. Relevant entry rate for the period analyzed is 9%, taken from Akcigit and Ates [2023]. Entry cost parameter is pinned down by matching the entry rate with the data. Finally overall level of $\bar{\alpha}_i$ is recovered by matching the model implied growth rate and its data counterpart which is taken to be 1.37% from Akcigit and Ates [2023]. High absolute level of research productivity increases the frequency with which inventors come up with new ideas, thus increasing the growth rate of aggregate productivity $\mathcal{A}(t)$.

1.5 Results

In this section, I present estimation results and resulted optimal place-based R&D policies in two stages. In the first stage, I assume that the US economy is comprised only of the top 10 states in terms of patenting. The reasons I focus on these states are three folds. Firstly, it is easier to discuss estimation results and the resulted optimal policy with fewer locations. Secondly, these states are well-known as being innovation locomotives of the US. Lastly,

Table 1.3: Top 10 states in patenting

State code	**State name**	**Patent share in 2005**
CA	California	25%
TX	Texas	7%
NY	New York	6%
MA	Massachusetts	5%
WA	Washington	4%
MI	Michigan	4%
IL	Illinois	4%
NJ	New Jersey	4%
MN	Minnesota	4%
PA	Pennsylvania	3%
Total		65%

patent citation flows are the most intensive among these regions. In order to draw conclusions on the effect of knowledge network on parameter estimates, I estimate two versions of the model. The first one is performed assuming $\varphi = 0$, i.e. the knowledge spillovers across regions are shut down. The second estimation is performed for the baseline model in which knowledge spillovers are active, $\varphi = 0.5$. After comparing estimation results, I proceed with the implied optimal policy for both estimations. Finally, I do several counterfactual exercises in order to assess the importance of knowledge spillovers and amenities for the characterization of optimal policy. In the second stage, I perform the estimation for the whole US economy, i.e. 51 states including DC. Qualitatively similar effects of knowledge spillovers arise in the full estimation as in the case for top 10 states. I, then, solve for the optimal R&D policy and discuss welfare implications.

1.5.1 Results for top 10 patenting states

In this section, it is assumed that the US economy is comprised of only ten states that produced the most of patents in 2005. Total share of these states in aggregate patenting is 65%. California (CA) comes first with a share of 25% followed by Texas (TX) with 7% and

Table 1.4: Parameter estimates for top 10 states - Without knowledge spillovers

States		Parameters		Allocations	
Code	**Name**	**Prod.** $\bar{\alpha}$	**Amenity** A	**Inventors** γ^R	**Workers** γ^L
WA	Washington	0.48	0.42	0.08	0.04
MA	Massachusetts	0.42	0.48	0.08	0.06
CA	California	0.42	1.00	0.35	0.24
MN	Minnesota	0.38	0.43	0.05	0.04
MI	Michigan	0.34	0.54	0.07	0.07
NJ	New Jersey	0.32	0.52	0.06	0.07
NY	New York	0.29	0.75	0.09	0.14
TX	Texas	0.29	0.79	0.10	0.15
IL	Illinois	0.28	0.63	0.06	0.10
PA	Pennsylvania	0.27	0.62	0.06	0.09

Note: Rows are ordered from highest $\bar{\alpha}$ to lowest

New York (NY) with 6%. The smallest shares belong to New Jersey (NJ), Minnesota (MN) and Pennsylvania (PA) with respective shares of 4%, 4%, and 3%. Table 1.3 illustrates the huge concentration of patenting even among the top ten most innovative locations.[12]

Table 1.4 illustrates parameter estimates for the sample states. In this version of the estimation, the knowledge network is inactive, i.e. $\varphi = 0$. Endogenous research productivity α_i is exactly equal to location fundamentals measured by $\bar{\alpha}_i$. The most research productive state is estimated to be Washington with a research productivity of 1.8 times that of the least productive state Pennsylvania. In terms of amenities, California has the highest, while Washington has the least, less than half of California. Although California has the highest fraction of inventors among these states, it also has the largest employment share with 24%. For the identification of research productivities across locations, number of inventors alone is not informative as inventors also value location amenities. As discussed in Section 1.4.2, the ratio of inventor share to worker share is the moment that identifies research

12. Worker and inventor shares are recalculated among ten states. For instance, California, the state with the highest share of inventors, is home to 22% of all inventors in the US, while, among the top 10 states, its share increases to 35%. Other two aggregate targets, growth and entry rates, are kept same in their original values, 1.37% and 9%, respectively.

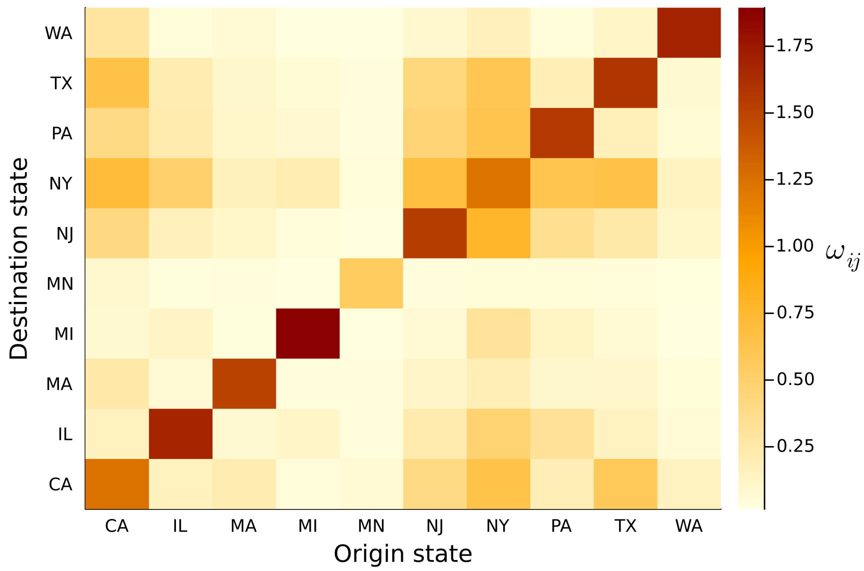


Figure 1.7: Estimated knowledge network Ω for ten states

productivities. As an example, the most research productive state Washington is home to 8% of inventors with an employment share of 4%. The ratio of the two is higher than that for California. Another example is the second largest state in terms of inventor count, Texas. 10% of inventors among ten states are observed to locate in Texas, while 15% of employment takes place in there. Thus, inventors choose Texas relatively less frequently than workers, informing the model inversion about relatively lower research productivity in Texas. On the contrary, amenities are directly identified by worker shares putting Texas on the second place in terms of relative amenities.

When knowledge spillovers are allowed between locations, the ranking of states in terms of research productivities change significantly. In order to proceed with model inversion, first knowledge diffusion rates ω_{ij} are estimated from the patent citation (1.42). Figure 1.7 depicts the estimated network matrix Ω in a heatmap plot. In this figure, origin states, represented as columns, refer to states from where ideas diffuse to the rest. Destination states, represented as rows, are the states to where ideas diffuse from origins. Some observations are in order. Firstly, diagonal terms have the highest values suggesting that within location

Table 1.5: Parameter estimates for top 10 states - With knowledge spillovers

States		**Parameters**	
Code	**Name**	**Prod.** $\bar{\alpha}$	**Amenity** A
WA	Washington	0.18	0.42
MN	Minnesota	0.17	0.43
MA	Massachusetts	0.14	0.48
MI	Michigan	0.12	0.54
CA	California	0.11	1.00
NJ	New Jersey	0.07	0.52
IL	Illinois	0.06	0.63
TX	Texas	0.06	0.79
NY	New York	0.05	0.75
PA	Pennsylvania	0.05	0.62

Note: Rows are ordered from highest $\bar{\alpha}$ to lowest

spillovers are stronger than spillovers across different states, in line with the findings of Jaffe et al. [1993]. Secondly, knowledge network Ω is observationally a symmetric matrix. For instance, CA exports knowledge mostly to WA, TX, PA, NY, and NJ (CA column), and it also imports knowledge mostly from these states (CA row). Thus, we can conclude that if i is connected to j, it is likely that j is also connected to i. Connections between states are mostly bilateral. Thirdly, as can be inferred from columns, California, New York, and New Jersey are the most upstream states in the flow of ideas. That is, these states export ideas relatively faster than other states. Lastly, Minnesota (MN) stands alone as being the least connected state both in terms of idea exports and imports.

Table 1.5 shows parameter estimates based on estimated Ω matrix. Amenity estimates are same as before, however, inclusion of knowledge network to estimation alters the estimates for $\bar{\alpha}_i$. In particular, Minnesota rises to second place in terms of research productivity while California declines to fifth place. The reason is that observed number of inventors in Minnesota (relative to its workers) can only be rationalized with a high $\bar{\alpha}_i$ estimate as Minnesota stands alone as the state that benefits the least from knowledge spillovers. Similarly, overall research productivity α_i in California mostly stems from network effects so that in terms of exogenous research productivity $\bar{\alpha}_i$, California declines from third to fifth place.

Optimal policy. The change in estimated allocation of exogenous research productivity across locations has implications on the optimal policy. In an environment without knowledge spillovers between locations, optimal policy only corrects the dispersion in idiosyncratic location preferences of inventors. In decentralized equilibrium without spillovers, there is always positive measure of inventors who idiosyncratically value the least research productive state the most among all, although they earn a very low wage there. They compensate the low productivity (and resulted low earnings and consumption) with their private value for the location. However, this is not aligned with the objectives of the planner, as the planner also cares about the effect of inventors on the rate of economic growth. Thus, optimal policy aims to relocate researchers towards the most research productive states by place-based R&D subsidies. Relocation of all of the inventors to the most productive state is extremely costly in terms of the forgone consumption due to taxation, as Frechet taste distribution has heavy tails. The trade-off that the planner faces is the tension between higher consumption in the future due to higher output growth, and lower current consumption due to taxation.

Figure 1.8 shows the inventor allocation under optimal policy for the case without knowledge spillovers (maroon bars) compared to the observed allocation in the data (gray bars). In order to achieve the optimal allocation, the planner subsidizes R&D expenditures only in four states: WA by 41%, MA by 31%, CA by 30%, and finally MN by 19%. This policy is financed by a permanent 1.5% uniform labor income tax. Under optimal policy, inventor allocation is more concentrated towards the most three productive states, WA, MA and CA. An interesting case is Minnesota (MN). Although it is subsidized, Minnesota experiences a decline in its inventor share under the optimal policy. That is, inventor share of MN would have been lower without subsidies. Optimal policy causes a welfare increase of 0.47% in consumption equivalent terms, while the long run growth rate of the economy rises from

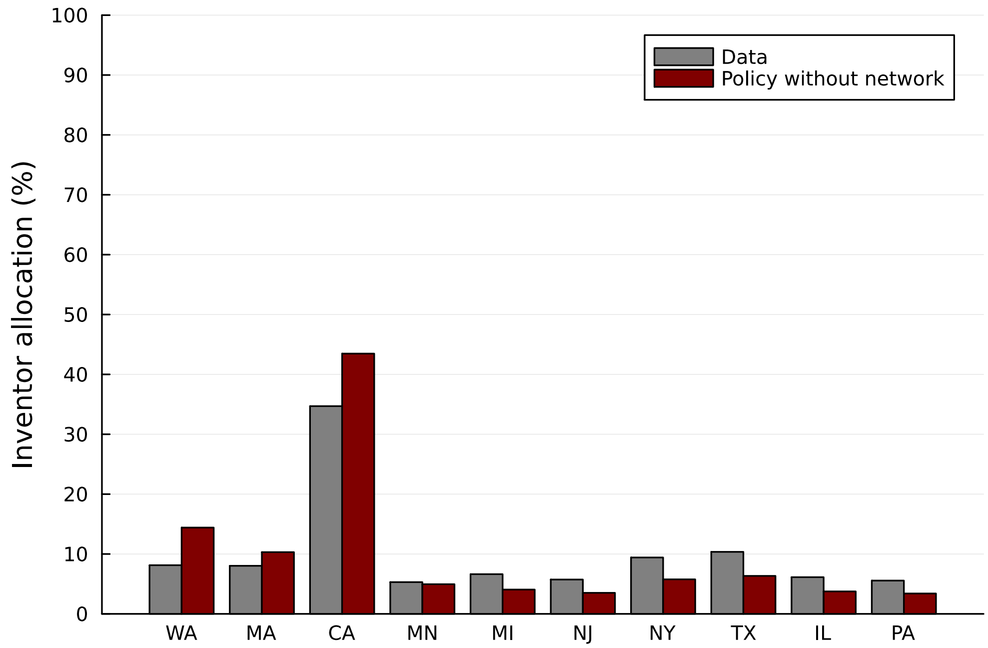


Figure 1.8: Inventor allocation under optimal policy - Without knowledge spillovers

Note: States are ordered left-to-right from the highest $\bar{\alpha}$ estimate to the lowest under "without" spillovers estimation.

1.37% to 1.41%. This exercise verifies the main intuition behind the workings of the optimal policy. Although there are no knowledge spillovers in this version of the model, the planner still wants to correct for the dispersion in idiosyncratic location tastes across inventors, as explained above. As spillovers are absent, the welfare gain from the policy is moderate.

When knowledge spillovers are present, the optimal policy starts reacting to the linkages between locations. In the baseline model, inventors do not internalize their effect on the productivity of other inventors through knowledge spillovers. While original incentive of the planner to relocate inventors to most productive states is still operating, spatial linkages makes the policy nontrivial. The trade-off that the planner faces is between research productivity and network centrality of locations. It might not be optimal for the planner to relocate inventors to most productive states if those locations are not connected well with the rest of the geography. Instead, it might be a better strategy to relocate inventors to moderately productive states but with strong linkages to the rest, both in upstream and

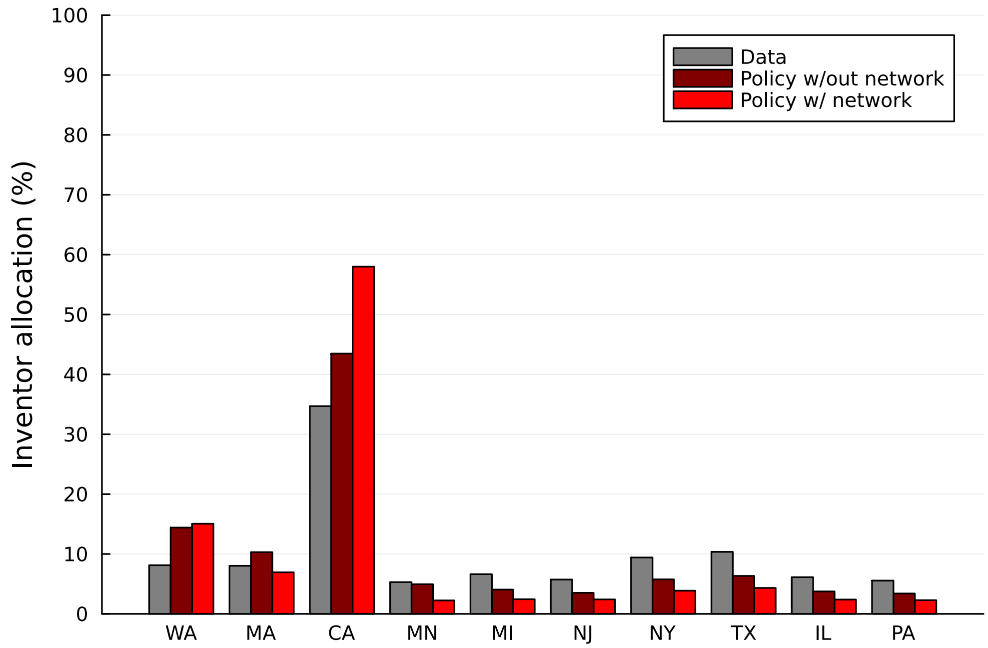


Figure 1.9: Inventor allocation under optimal policy - With knowledge spillovers

Note: States are ordered left-to-right from the highest $\bar{\alpha}$ estimate to the lowest under "without" spillovers estimation.

downstream sense. By this way, the planner maximizes the extent of knowledge spillovers. It should be also noted that research productivity and growth considerations are not only factors that shape the optimal policy. Relative amenities directly affect the social welfare function (1.40). The effect of amenities on optimal policy is through two channels. The first channel is the direct effect. All else equal, the planner wants to benefit from highest amenities in the country. The second channel is through the effect of amenities on the cost of the policy. If a location simultaneously have both high amenity and research productivity, then reallocation of inventors to that location would be less costly in terms of taxation, as inventors would be more likely to migrate to that location due to high amenities in there. However, if amenities and research productivities are not aligned well, then the planner has to subsidize R&D very heavily in order to be able to convince inventors to migrate there. This increases the taxes imposed, hance the cost of the policy.

In order to assess the discussed effects of the knowledge network on policy, place-based

R&D subsidies are solved for the baseline model with knowledge spillovers and compared to the previous case. That is, optimal policy is solved for the parameterization given by Table 1.5. Figure 1.9 shows the inventor allocation under the optimal policy in this case (depicted by red bars) while comparing it to previous policy and the data. In this case, subsidy rate in WA and CA rises to 49% in each, whereas it declines slightly to 30% for MA. MN is not subsidized anymore. Qualitatively, the planner stops allocating inventors only to the most productive states. The most clear example is Minnesota (MN). In the new policy, Minnesota experiences a stark decline in its share of inventors. The reason is that it is not well connected within the knowledge network (as suggested by the Ω matrix in Figure 1.7), so the social value of high research productivity in Minnesota decreases as other locations do not benefit much from the spillovers from Minnesota. Similarly, Massachusetts (MA) also experiences a decline in its inventor share. In the new policy, inventors are relocated towards WA and CA mostly from MA and MN.

Inventor concentration across locations under the new policy results in higher than before. In addition to the network effects discussed above, another reason for this result is that knowledge spillovers are the strongest within states (high diagonal elements of Ω). Therefore, inventors cause an agglomeration type of spillovers in their own locations, which calls for more concentration under optimal policy. Finally, welfare increase in a model with knowledge spillovers is found to be higher, i.e. 1.68% consumption equivalent increase in welfare with a growth rate of 1.49%. Although the size of welfare effects are not directly comparable between two models, intuitively, we can argue that knowledge spillovers and resulting increasing returns to scale makes the optimal policy more effective in terms of welfare increases. In order to draw more meaningful comparison between the two policies, i.e. one that respects the knowledge spillovers between locations, and the other that does not consider linkages, we can implement the first policy within the estimated model with knowledge network. This exercise results in a welfare increase of 1.20% in consumption equivalent terms, which is

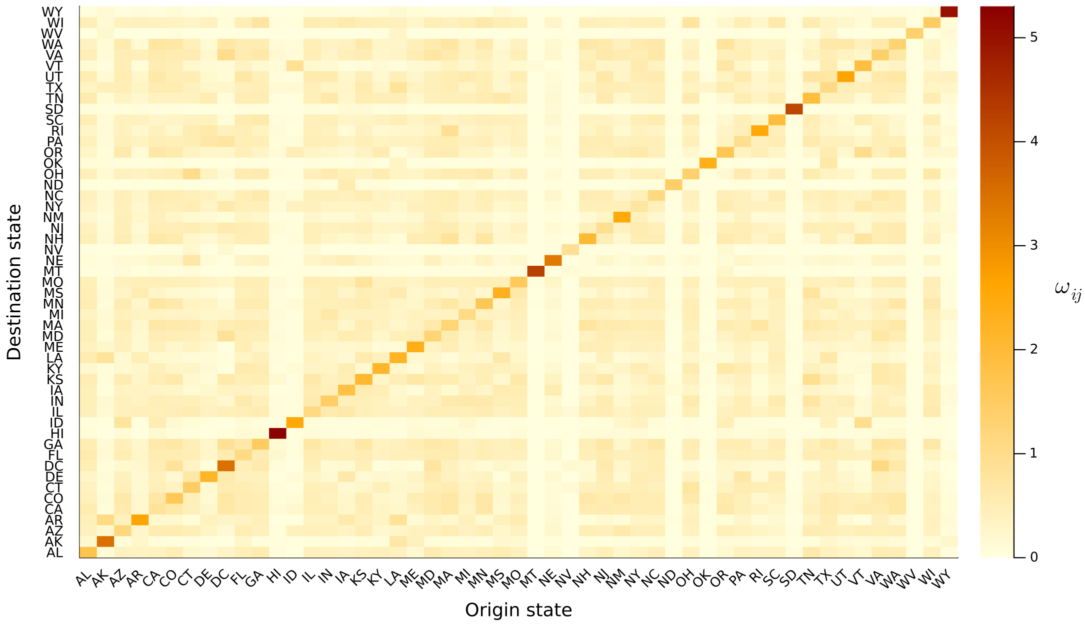


Figure 1.10: Estimated knowledge network Ω

0.48 percentage points lower than the welfare increase under the policy that respects the knowledge spillovers.

1.5.2 Results for all states

In this section, I present estimation results of the model with all the states in the US (51 states including DC). Then I compare model implied untargeted moments with the data in order to validate parameter estimates. Finally, I solve the optimal place-based R&D subsidy policy. My findings can be summarized as follows. Model fit to untargeted moments is good giving confidence on the model's validity. Optimal policy calls for concentration of inventors in a few states on both West and East coasts such as Washington, Massachusetts, California, and Vermont. Social welfare increases 1.8% in consumption equivalent terms as a result of the proposed place-based R&D policy. Increase in welfare is associated with a 0.14 percentage points increase in annual growth rate of the economy, from 1.37% to 1.41%.

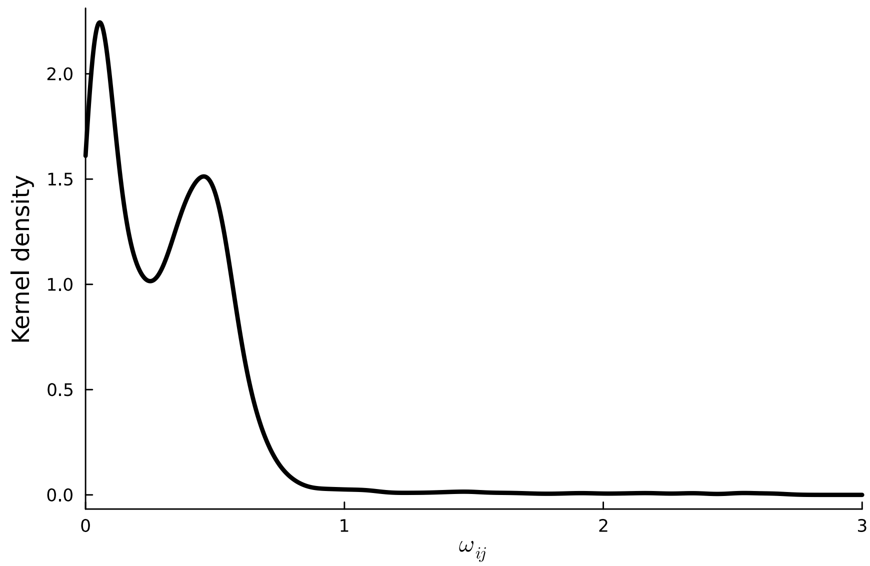


Figure 1.11: Kernel density of diffusion rate estimates

Estimated knowledge network Ω. The heatmap of estimated Ω matrix is given by Figure 1.10. Similar results are observed in the full estimation of the matrix. First of all, diagonal elements are considerably higher than off-diagonal elements suggesting strong within location spillovers. It can be argued that strength of connections between states are usually bilateral. Some states are isolated from the rest of the network both in upstream and downstream sense, such as Idaho, Montana, South Dakota, West Virginia. The (unweighted) mean of ω_{ij} is 0.31 suggesting a lag of 3.2 years in idea diffusion across states. The histogram of estimates for diffusion rates ω_{ij} is plotted in Figure 1.11 suggesting a bimodal distribution of pairwise diffusion rates across US states.

In order to test the validity of ω estimates, I regress estimated ω_{ij} on some observed characteristics of state pairs. These are pairwise physical distance between i and j, and academic citation shares, migration flows, number of air passengers, and trade flows, both from i to j, and from j to i. Only the coefficient estimates of physical distance and academic citation shares are significant, and their signs are as expected. In words, estimated ω_{ij} decreases in physical distance between i and j. Moreover, it is positively correlated with

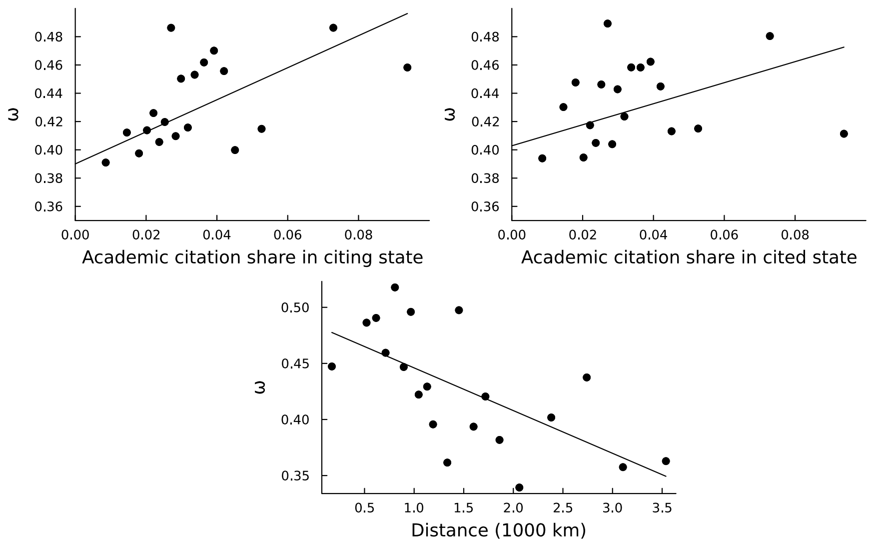
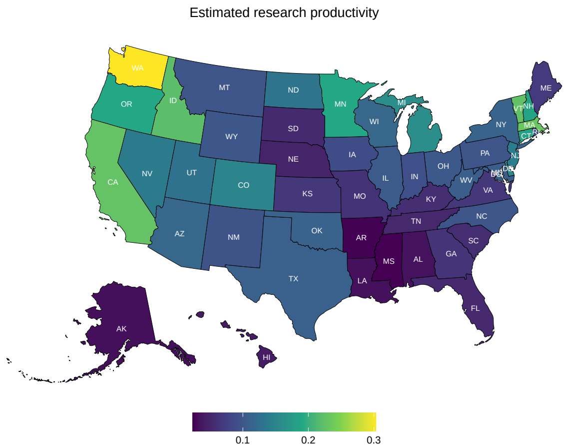


Figure 1.12: Estimates of ω_{ij} and their correlates

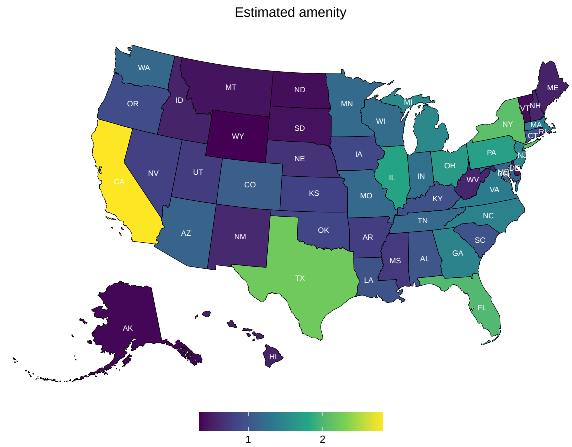
probability of academic papers published in i citing papers from j, and vice versa. Binscatter plots for the relationship between ω_{ij} estimates and observed characteristic are shown in Figure 1.12.

Location specific parameter estimates. Figure 1.13a illustrates the distribution of estimated research productivity $\bar{\alpha}$ across US states. The most research productive state is estimated to be Washington with a value of $\bar{\alpha} = 0.303$ followed by Massachusetts and California. The least research productive state is Mississippi with a value of $\bar{\alpha} = 0.024$. The mean of $\bar{\alpha}$ estimates is 0.110, and their standard deviation is 0.063. As is clear from Figure 1.13a, physical proximity is an important determinant of the spatial distribution of $\bar{\alpha}_i$. That is, closer states are also similar in terms of research productivity. West coast states represented by Washington, Oregon, California, and perhaps including Idaho are most productive states in research along with a clustering on the East coast represented by Massachusetts, Vermont, Connecticut, and New Hampshire. In the Midwest, Minnesota and Michigan stand out.

Amenities, on the other hand, do not seem to be correlated much with research produc-



(a) Research productivity $\bar{\alpha}_i$

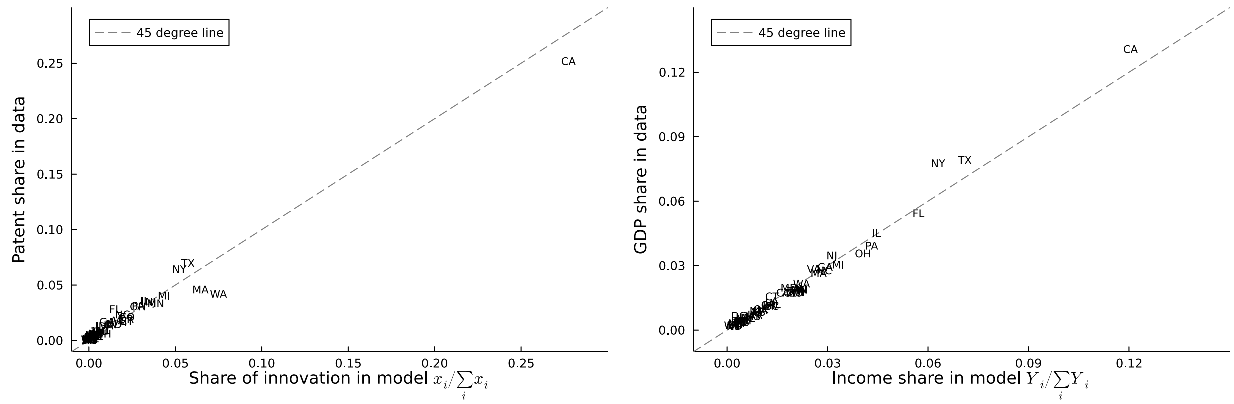


(b) Amenity A_i

Figure 1.13: Estimated location specific parameters

tivities. California has the highest amenity estimate followed by Florida, New York, Texas, Illinois.

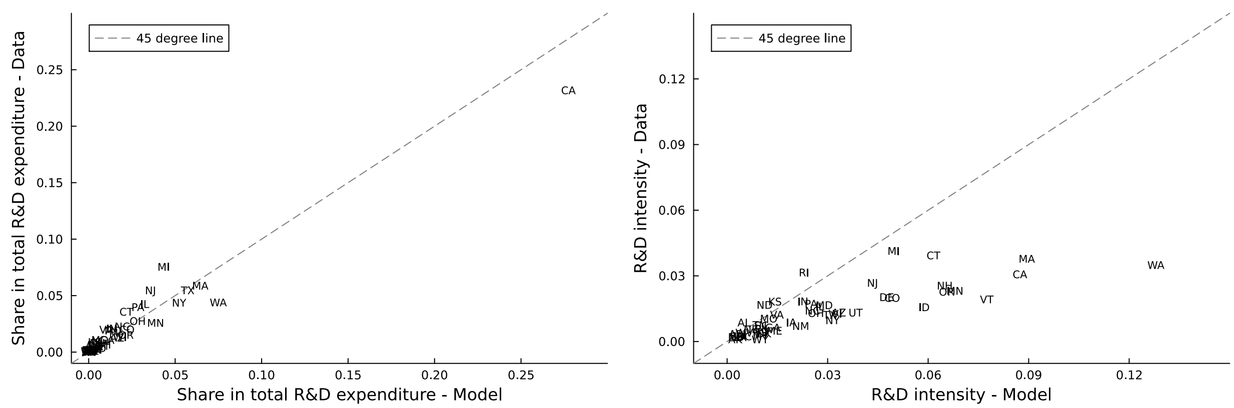
Model fit for untargeted moments. In this section, the model is tested in terms of its fit to untargeted moments such as relative patenting rates of states, GDP shares, share of states in total R&D expenditure, and finally, R&D intensity of states defined as the ratio of R&D expenditure to state's GDP. The model performs well along these dimensions giving confidence on the validity of the model. Rate of patenting in a state in the model is given by x_i. The share of state's patents produced in a given year in total number of patents produced in the US can be measured by $x_i / \sum_i x_i$. Figure 1.14a shows the comparison of this moment between model and the data. Most of the observations lies on the 45 degree line implying almost perfect match. Another untargeted moment is GDP share of states. The GDP of a state in the model economy is defined as the total income of agents (including profits) located in the state, $Y_i \equiv (1+d)\left[W_i^L(t)L_i + W_i^R(t)R_i\right]$. Share of state's GDP in total US GDP is given by Figure 1.14b showing a very good match between model and the data. States vary slightly around the 45 degree line.



(a) State's share in patenting

(b) Amenity A_i

Figure 1.14: Model's fit to untargeted moments



(a) State's share in total R&D expenditure

(b) State level R&D intensity

Figure 1.15: Model's fit to untargeted moments

Finally, I check implied state level R&D expenditures from the estimated model, and compare it to the data obtained from National Science Foundation's (NSF) National Patterns of R&D Resources for the year 2004. This data provides state level R&D expenditures of private industry and government. Only industry R&D expenditures are included in state level R&D spending. In the model, state level R&D expenditure is given as the total wage bill of researchers employed in both incumbent and entrant firms. It is equal to $W_i^R(t)R_i$, as R&D subsidy rates are taken to be zero in the benchmark estimation. Figure 1.15a compares the share of state's R&D expenditures in the model and in the data. Again, the model fit is very good in terms of relative R&D spending across states. On the right panel, state

Table 1.6: Subsidy rates under optimal policy

State code	State name	Subsidy s
WA	Washington	0.57
MA	Massachusetts	0.48
CA	California	0.48
VT	Vermont	0.39
MN	Minnesota	0.35
OR	Oregon	0.34
NH	New Hampshire	0.32
CT	Connecticut	0.24
CO	Colorado	0.09

level R&D intensities are plotted. R&D intensity of a state is defined as the ratio of R&D expenditures to state level GDP. The model predicts a higher level of R&D intensity for most of the states relative to the data. However, model implied moment and data counterpart are positively correlated.

1.5.3 Optimal policy

Under the optimal place based R&D subsidy scheme, only 9 states receive R&D subsidies. Most heavily subsidized state is Washington with 57%, and the least subsidized state is Colorado with 9%. Remaining states do not receive R&D subsidies. Table 1.6 lists the R&D subsidy rates across states from the highest to the lowest. On the West coast, neighbor states California, Oregon and Washington receive R&D subsidies as this region of the country is the most R&D productive, and they are relatively connected with the rest of the states. Minnesota is the only state that is subsidized in the Midwest, while the small region around Massachusetts benefit from subsidies as well. In terms of inventor allocation under the optimal policy, only seven states out of nine increase their inventor share. These two states that experience a reduction in inventors are Connecticut and Colorado. The reason is that concentration of inventors under the optimal policy is strongly towards the most research productive states, but the planner does not want to relocate much from Connecticut and

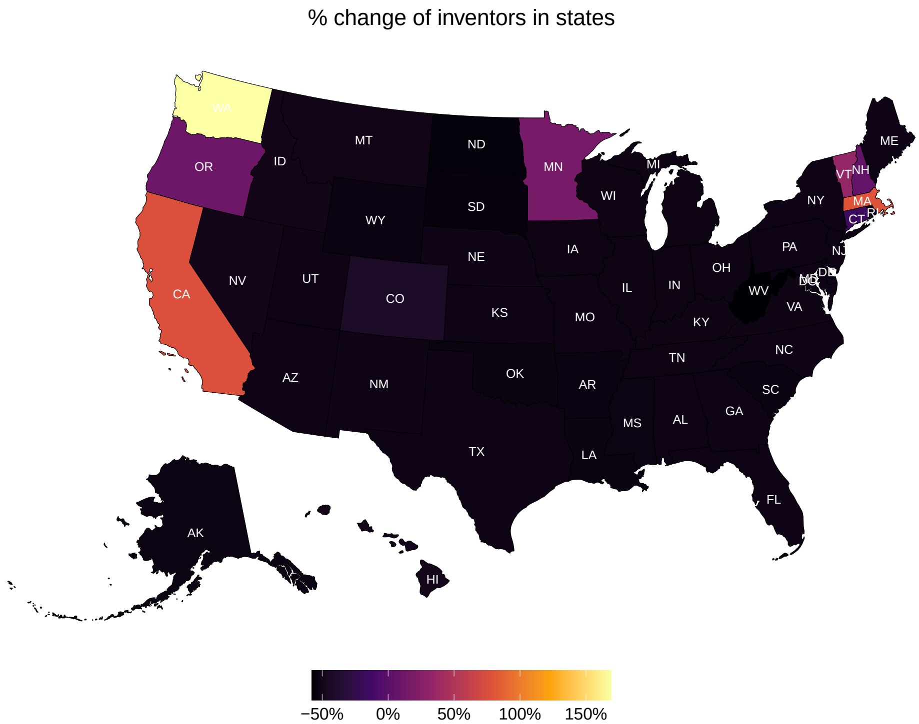


Figure 1.16: Percent change in number inventors under policy

Colorado. By R&D subsidies the decline in these two states is mitigated.

The optimal policy calls for concentration of inventors in top productive states such as Washington, Massachusetts, and California. The percent change in number of inventors between the optimal policy and the data for each state is depicted in Figure 1.16. Except the top five states, the rest lose almost half of their inventors under the new allocation. Washington experiences a 169% increase in its inventors relative to the data. Its share rises from 5.2% to 14%. Massachusetts and California observe similar increase in their inventors, 79% and 77%, respectively. Under the new allocation, California still has the highest share of inventors with 39.2%. Among the losing states, small states lose the most. Table 1.7 shows the top 5 gaining and losing states in terms of inventor counts under the optimal policy.

The reason for increased concentration of inventors under the policy is close geographical connections between states and strong within state knowledge spillovers. Although there

Table 1.7: Top gaining and losing states

State code	**State name**	γ^R	$\gamma^{R\star}$	$100\frac{R^\star - R}{R}$
	Top 5 gaining states			
WA	Washington	5.2	14.0	+169%
MA	Massachusetts	5.1	9.2	+79%
CA	California	22.2	39.2	+77%
VT	Vermont	0.40	0.53	+33%
MN	Minnesota	3.4	3.9	+16%
	Top 5 losing states			
WV	West Virginia	0.14	0.06	-58%
ND	North Dakota	0.11	0.05	-56%
SD	South Dakota	0.06	0.03	-55%
LA	Louisiana	0.33	0.15	-54%
OK	Oklahoma	0.49	0.22	-54%

is no reduced form agglomeration spillovers in the model, high within diffusion rates ω_{ii} implies local inventors benefit most from the local spillovers. This induces the planner to relocate more to the most productive states. However, geographical proximity also seems an important determinant of new inventor allocation as diffusion rates and proximity are strongly correlated as shown previously.

The overall welfare increases 1.8% in consumption equivalent terms. This welfare increase is achieved even without allocating the labor force towards being researchers, rather it is due to geographical reallocation of a constant pool of inventors in the country. This policy exercise shows the importance of knowledge spillovers even inside a country between localities, and points to a significant level of welfare loss due to imperfect knowledge spillovers specific to the innovation process. The 1.8% increase in welfare is associated with a 0.14 percentage points increase in the growth rate of the economy. As inventors benefit more from knowledge spillovers, their research productivity rises helping them innovate more frequently. The growth rate rises from the targeted value 1.37% to 1.51% under the optimal policy.

1.6 Conclusion

In this paper, I study the spatial allocation of inventors across US states and the effect of their location choice on the aggregate rate of innovation and growth, with a particular focus on knowledge spillovers across states. Empirically, it is shown that innovative activity is spatially concentrated more than other indicators such as employment and GDP. Furthermore, distribution of inventor to worker ratio across US states is highly right skewed suggesting that inventors prefer certain locations more than other workers. These locations coincide with the innovation hubs of the US such as California, Washington and the Northeast corridor. Finally, I show the extent of spatial concentration in patent citations, and argue that the variation in patent citation lags across citing-cited state pairs is particularly informative on the extent of knowledge linkages between them. These spillovers are intertemporal in the sense that future inventions benefit from the old ideas previously invented, which reflects itself as citations between patent documents.

On the theory side, a novel endogenous growth model is built with inventor and worker migrations in space, and mobile entrants who create firms in which inventors are employed for R&D purposes. The model is equipped with a knowledge diffusion network between locations which is estimated from patent citations. Inventors in the model do not internalize the effect of their location choice on the diffusion of ideas to other locations. Thus, the planner corrects for this externality by place-based R&D subsidies while taking into account heterogeneous linkages between US states.

Location specific parameters, amenities and exogenous research productivities, are recovered with a model guided inversion procedure which exactly matches observed worker and inventor allocations across US states. The unknown knowledge diffusion network is estimated from a patent citation equation that is derived from the model. It is shown that states that are close in distance are more likely to be connected, academic citation flows are positively correlated with the strength of connections, and within-location spillover rates are the highest. Based on all parameter estimates, the optimal placed based R&D subsidy policy is found as to maximize the social welfare function. The policy calls for even more spatial concentration of inventors although the model is absent from standard within-location agglomeration spillovers. Moreover, the policy respects the flow of ideas in space in the design of the place-based subsidy rates. The optimal policy increases the social welfare by 1.8% in consumption equivalent terms. The increase in welfare is associated with a 0.14 percentage points increase in the aggregate growth rate of the economy due to the maximum utilization of the knowledge spillover network between states as a result of the policy.

CHAPTER 2

READY, SET, UNIFY: THE UNEVEN RACE BETWEEN TRABANTS AND BMWS

This chapter is coauthored with Ufuk Akcigit, Sina Ates, Matthias Mertens and Steffen Mueller.

2.1 Introduction

Even 30 years after the reunification, regions in the former East Germany (GDR) live in considerably different economic conditions compared to their West German counterparts. While there have been notable improvements in the living standards of Eastern households, as of 2021, GDP per worker in the East still remains about 20 percent lower than that of West Germany (Figure 2.1), with a comparable gap in the average wage level, despite massive financial support provided to the East via various schemes to reinforce its economic development (Figure 2.2). The enduring economic gap between East and West Germany continues to receive perennial attention from both the general public and academia.[1] It also remains a salient topic in political election campaigns and contributes to political polarization in Germany and the ascent of populist parties (see, e.g., Weisskircher [2020]; Politico [2021]; The Economist [2023]).

In this study, we investigate the causes of the persistent regional disparities in productivity, their link to the innovation incentives of regional firms in the unified economy, and the productivity and welfare implications of policies (financed by the West to support Eastern development) that could potentially support the economic convergence of the two regions. We develop a dynamic general equilibrium model of firm-level competition, innovation, and

1. See, for instance, The Economist [1999], The Economist [2014a], The Economist [2014b], Euractiv [2023], Becker et al. [2020], and Burda [2020].

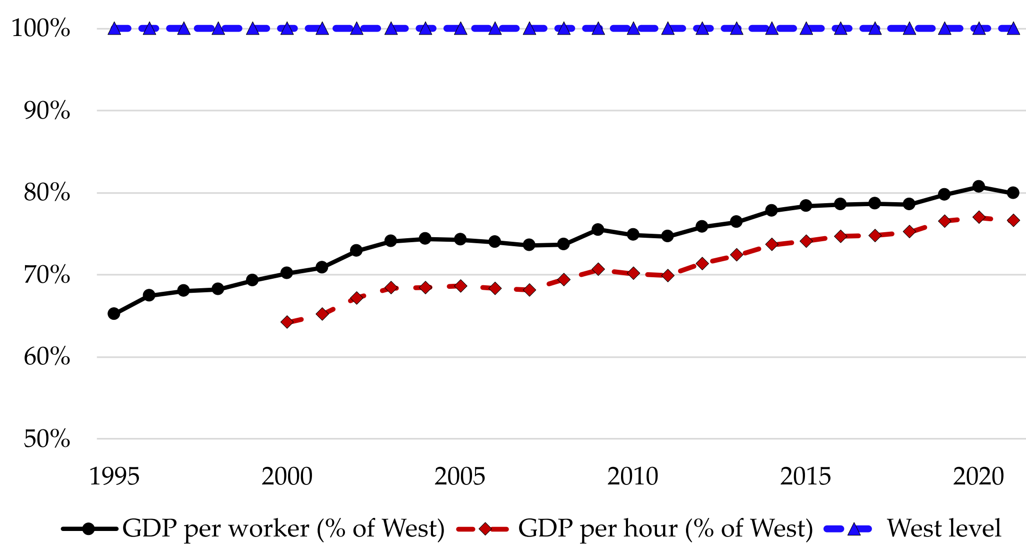


Figure 2.1: GDP per worker and wages, East in percent of West

Notes: The y-axis displays the level in terms of percent relative to the West. Source: National Accounts of Germany.

endogenous growth with two regions that allows us to examine the transitional dynamics set off by the unification of the regional economies. We combine it with comprehensive micro-level data; in particular, information on firms' prices in very narrow product classes plays an instrumental role in determining the evolution of product qualities and, thus, the relative technological performance of firms in the East and West. Our results highlight that (i) the inferior technological level of Eastern firms initially discouraged them from competing with Western rivals, reducing their incentives to improve their technology, (ii) existing policies did not focus on support to technological upgrading of Eastern firms, and (iii) policies that focus on technology transfers to the East—via direct licensing or spillovers—prove more efficient than financing direct support to Eastern firms' R&D.

Our analysis is motivated by the well-known observation that the East lacked competitive products and viable firms vis-à-vis the West at the time of reunification (Akerlof et al. [1991]; Collier [1991]). Figure 2.3 uses the example of the Trabant (East) and BMW (West) to illustrate the profound difference in product quality and production technology between

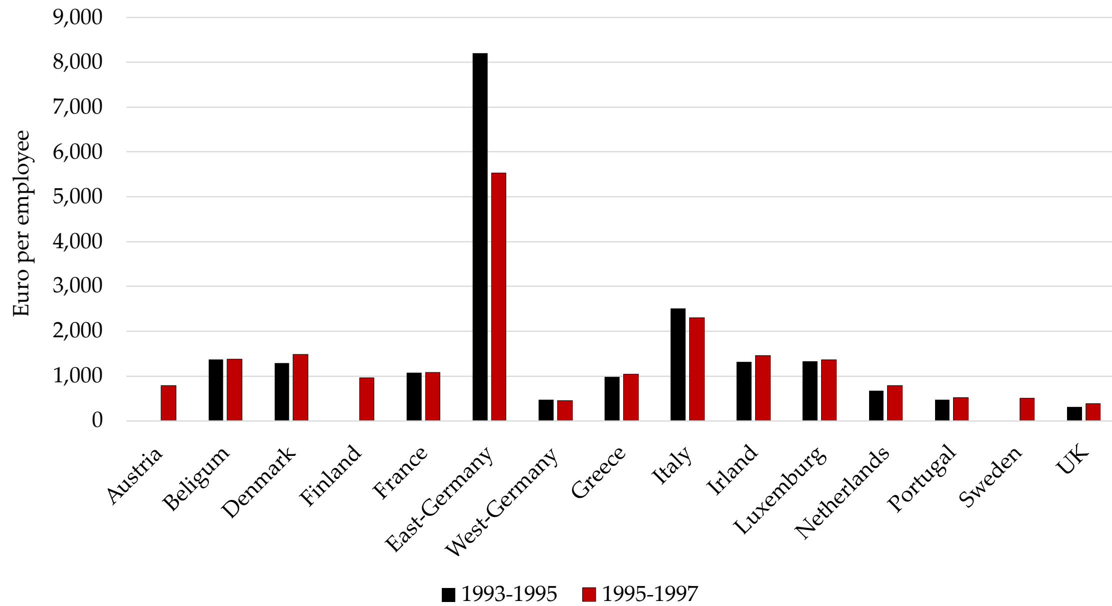


Figure 2.2: Manufacturing subsidies per employee, by countries

Notes: Manufacturing subsidies in Euro per employee, by countries. Values of 1996. Source: BMF (1999).

the former GDR and West Germany.[2] Figure 2.3a displays the Trabant 601, produced by Sachsenring Automobilwerke Zwickau from 1964 to 1990. Following the German reunification, the Trabant faced new major Western car manufacturers such as Mercedes, Audi, Volkswagen, and BMW, as new competitors. Figure 2.3b shows a BMW M3 cabriolet, which was produced from 1986 to 1991 by BMW. Even after substantial price cuts, Trabant could not retain a reasonable share of the market, and its production was discontinued swiftly.

The enormous technological disparity between Eastern and Western producers contributed to the massive shake-up of the industry in the East in the wake of the reunification, triggering an unprecedented contraction on impact. Still, helped by massive financial support from the West, the Eastern economy was able to recover swiftly, with its GDP per worker reaching

2. The automobile industry is just one example of numerous industries where Eastern firms faced a significant quality disadvantage compared to their Western counterparts. Upon entering world markets, Eastern firms had to contend with well-established Western brands, such as Adidas, Puma, Hugo Boss (clothing), Nivea (cosmetics), Haribo, Rittersport (food), Bosch, Braun (electronics), and Quelle (retail). Some of these brands even dominated their respective sectors globally.

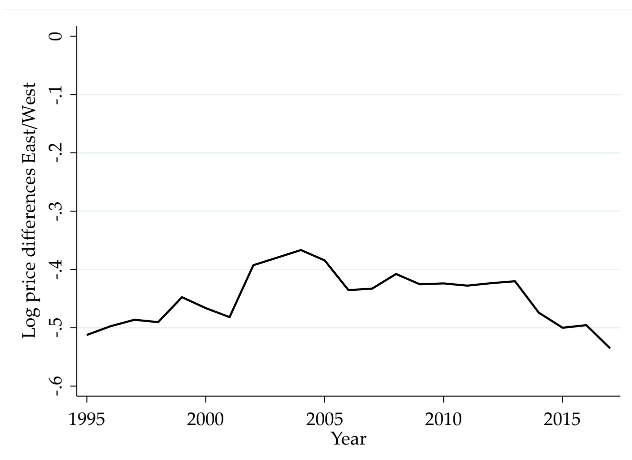
(a) Trabant 601, 1964-1990

(b) BMW M3 Cabriolet, 1986-1991

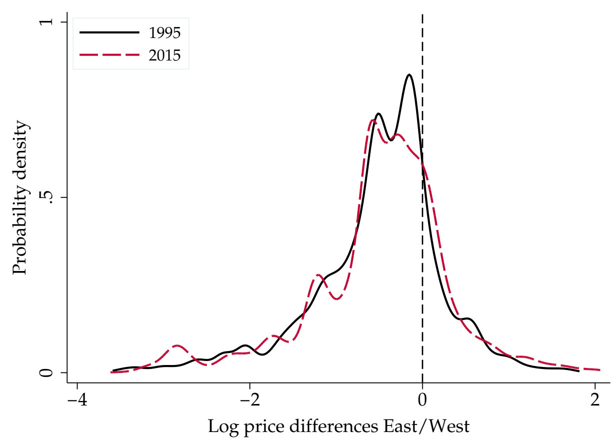
Figure 2.3: Trabant vs. BMW

Source: Trabant: https://imagebee.org/vehicles/trabant-601/, accessed July 11th 2023. BMW: https://www.autoevolution.com/news/the-bmw-m3-e30-convertible-was-the-embodiment-of-1980s-open-top-madness-192380.html, accessed July 11th 2023.

two thirds of the West German level by the mid-1990s. However, the convergence has been remarkably slow since then (Figure 2.1). Importantly, most East German producers that remained in business continued to produce goods that are of inferior quality and compete through lower prices, only taking advantage of lower production costs in the East (Mertens and Müller [2022]). Figure 2.4 showcases the striking persistence of the price difference be-



(a) Aggregate price differences



(b) Price distribution

Figure 2.4: Relative price differences, East/West

Notes: Panel A shows aggregate log output price differences between East and West German firms. Regional output prices are derived for each product as a sales weighted average. We aggregated across products using again sales-weights. Panel B shows the distribution of log product price differences between the East and the West. A negative value indicates lower prices in the East. German firm-product-level manufacturing sector data. 1995-2017.

tween Eastern and Western products, which can be observed even at the narrowly defined product level (Figure 2.4b)—a key observation that guides the theory that underlies our model.

The first goal of our analysis is to investigate the factors that led to persistent technological disparities between producers in the East and West and they ways in which they may have impaired economic convergence between the two regions. To this end, we build a general-equilibrium framework of endogenous firm dynamics based on step-by-step innovation models (Aghion et al. [2001]; Aghion et al. [2005]; Acemoglu and Akcigit [2012]; Akcigit and Ates [2021]). We extend the basic framework to include two regions that trade goods without frictions in a unified economy. The unified economy consists of a unit measure of tradable products—in addition to a non-tradable good. In each product line, one firm from each region produces a different variety, and these firms engage in Bertrand competition for higher market share. Their varieties differ in product quality, which affect their prices and the market share of the firm. Firms can enhance their product quality through successive innovations investing in R&D, while the firm with the inferior quality also benefits from technology spillovers from the superior firm improving their quality at an exogenous rate. Firms differ also in their production costs due to region-specific labor productivity, which influences their competitiveness in the marketplace and evolves endogenously to reflect the changes in relative wages across the two regions observed in the data. In addition, region-specific R&D efficiency drives comparative advantage in innovation.

Linking competing firms' relative technology level to their position in market competition and their share of the market, which in turn influence the innovation incentives of firms, the step-by-step innovation structure underlying our framework provides a natural setting to investigate how technological differences among firms interact with their innovation investments and determine endogenous dynamics of aggregate productivity growth and economic convergence. A salient implication of this setting, which is the focus of most studies of this framework, is the so-called escape-competition effect—that is, close market competition between firms intensify their innovation effort with the goal of getting ahead of their rival and escape competition. In our work, instead, a related but different force is key: the discouragement effect. This effect implies that firms that are technologically too laggard lose their hope of catching up with and outpacing their rival, which, in turn, depresses their incentives to innovate. As such, the discouragement of laggard firms perpetuates the technological disparities, weighing on regional convergence.

While the discouragement effect is a theoretical force that could have possibly hampered convergence, we quantitatively assess the plausibility of this relationship in a calibrated version of the model carefully disciplined by micro-level data. We calibrate the transitional path of the model to mimic the dynamics of the Eastern and Western German regions over the period between 1995 and 2015, including the patterns of relative income and wages. While time-varying specifications of regional labor force matched to the data capture the variation in regional population as a result of migration flows, an exogenous path of regional labor productivity fit to its empirical counterpart helps the model capture differences in production costs.

Most notably, our calibration strategy takes advantage of the detailed price data on the products of Eastern and Western German firms, which cover almost 6,000 product codes within the manufacturing sector. We use the distribution of relative average prices across the products produced in the two regions within each of these finely defined products—as displayed in Figure 2.4b—to discipline the model counterpart of relative price distribution across product lines and its evolution. The ultimate goal of this exercise is to inform the technology dynamics across the regions in the model, which admits a clear mapping between firms' relative technologies and their relative prices. As a result, the calibration of the price distribution disciplines the distribution of relative product qualities in the model, once the production costs are controlled. Hence, the detailed product-level price data proves instrumental in deriving the dynamics of product qualities through the lens of the economic theory. The calibrated model indeed points to persistent technology gaps between regions owing to the discouragement effect.

The calibrated model serves as a laboratory to examine several counterfactual policy settings. To start, we evaluate the implications of delaying reunification—an alternative debated fiercely at the time—which we model as minimal goods trade between the regions via high tariff barriers approximating autarky. The temporary increase in trade barriers depresses the technological progress—measured by the average quality of products in a region—and reduces welfare (in consumption-equivalent terms), because firms lose access to the market in the other region, and the reduced rents weigh on innovation incentives. Subsidizing concurrently R&D in the East could boost innovation effort, but the resulting benefits could offset the associated costs only over the longer policy horizons.

Next, we evaluate various transfer schemes to the East funded by the West. Two prominent schemes in the aftermath of the reunification constituted effectively lumps-sum transfers to households and firms, which we model as consumption subsidies to Eastern consumers and production subsidies to Eastern firms. Neither of these policies prop up the technological development of Eastern firms. They do improve welfare in the East but at the expense of the welfare of Western consumers. We find that channelling the same amount of resources to subsidize R&D expenditures by Eastern firms does indeed accelerates average product quality upgrading in the East. However, it leads to notable welfare losses in both regions over the short to medium policy horizons stemming from lump-sum taxation of Western consumers to fund the subsidies and the shift of production resources to R&D activity in the East, with benefits from higher product qualities accumulating and dominating welfare over time.

Given that doing R&D is less costly for Western firms in the calibrated economy, an alternative policy could be subsidizing R&D by Western firms, with the expectation that knowledge spillovers help lift average quality of Eastern products. This policy indeed proves more effective: average product quality growth picks up in the East, although, in relative terms, it falls behind the West faster than in the simulated baseline (calibrated) economy. Welfare increases in both regions (except for the immediate term in the West) and more so over longer horizons. Alternatively, we consider licensing of Western technologies to Eastern firms, with resources for transfers in other policy schemes analyzed being used to make up for the losses of Western firms that stem from reduced market share as Eastern firms improve their quality through licensing. A salient finding is that this policy front loads welfare gains in both regions as opposed to R&D subsidies. Therefore, a policy mix that smooths out welfare gains over time would be the combination of technology licensing by Western firms while subsidizing their R&D expenditures.

To sum, our analysis points to a novel technological factor coupled with the dynamics of market competition as the root cause of persistent product quality disparities between the East and the West and the slow pace of economic convergence. Policies that aim at boosting technology upgrading of Eastern firms could alleviate this problem. That said, our results emphasize that standard support schemes such as R&D subsidies to Eastern firms would not be welfare-improving over the relevant policy horizons in light of the estimated high cost of doing R&D in the East. More effective use of resources would need to rely on technology transfers from Western firms via licensing while supporting their innovative activity, respecting their comparative advantage and ensuring the flow of knowledge spillovers.

2.1.1 Literature Review

Applying a theory based on the step-by-step innovation framework, our study contributes to the burgeoning work on the quantitative analysis of this framework (Akcigit and Ates [2023]; Cavenaile et al. [2023]; Chikis et al. [2021]; Liu et al. [2022]). While not an open-economy model per se, our setting analyzes the integration of two regional economies, whose firms are trading and competing with each other, taking it closer to the recent open-economy examples of the step-by-step innovation framework (Akcigit et al. [2018]; Akcigit et al. [2024]; Choi and Shim [2024]). As in rare examples, the analysis concerns the transitional dynamics; yet, our model offers a richer and more nuanced environment of firm-level competition guided by the empirical analysis. It features regional firms that differ along multiple margins (product quality, productive efficiency, and labor cost) all of which evolve over time and affect firms' competitiveness, pricing decisions, and market share.

With regional economic disparities at its center, our study relates to the literature on regional convergence (Barro and Sala-i Martin [1992]; Blanchard and Katz [1992]; Sala-i-Martin [1996]; Caselli and Coleman II [2001]; Barro [2012]; Gennaioli et al. [2014]). The extraordinary experiment of German reunification has attracted notable attention in this context (Canova and Ravn [2000]; Burda and Hunt [2001]; Sinn [2002]; Burda [2006]; Uhlig [2008]). Most studies of the dynamics of economic differences between East and West Germany approach the issue with a focus on factor markets, particularly the labor market (Burda and Hunt [2001]; Hunt [2006]; Uhlig [2006]; Snower and Merkl [2006]; Fuchs-Schündeln and Izem [2012]; Findeisen et al. [2021]; Heise and Porzio [2022]).[3] Our study offers a complementary yet different take. It highlights another factor, which has been barely explored methodologically—namely, the technological underpinnings persistent regional economic disparities at the firm level. We uncover these technological factors and the endogenous mechanisms underlying them with the help of a structural dynamic general equilibrium model that is carefully disciplined by macro and detailed micro-level data. In this regard, our study sheds fresh light on the dynamics of regional economic development and convergence, which can face significant challenges even when it concerns two regions that are much alike and integrate under felicitous circumstances (Burda and Hunt [2001]).[4]

3. On a different note, Burchardi and Hassan [2013] examine the effect of social ties on regional growth dynamics.

4. For instance, East and West Germany do not differ much in salient characteristics such as climate,

Finally, our policy analysis contributes to the literature on industrial policy, which has received renewed interest recently (Aghion et al. [2015]; Acemoglu et al. [2018]; Akcigit et al. [2022]; Criscuolo et al. [2019]; Liu [2019]; Lane [2022]).[5] Evaluating policy implications over the transition of the economy and differentiating between the short and long term, our investigation shares the spirit of Akcigit et al. [2018] and Choi and Shim [2024]. Moreover, the distinction between R&D subsidies and licensing policies provides a more nuanced study of technology policy in a departure from recent applications (Atkeson and Burstein [2019]; Akcigit et al. [2018]; Akcigit et al. [2024]). In their seminal contribution, Acemoglu et al. [2006] distinguish between policies that foster technology adoption or innovation and provide a theoretical analysis of their implications depending on the development stage of a country. More recently, Choi and Shim [2024] provide a quantitative analysis of the role adoption and imitation policies played in South Korea's rapid economic development. Our work complements these papers by studying place-based policies, carefully evaluating their implications in the two regions and also contrasting them with various other policy alternatives.

2.2 Institutional Background

At the end of World War II, Germany was divided into a democratic capitalist West and a socialist East, a partition that lasted until the fall of the Berlin Wall on November 9, 1989, and the reunification in 1990. The integration of 16 million East Germans was a formidable challenge for the West. A key political issue in the 1990 federal election, which included participation from the East, was the timing and speed of economic integration. The central point of contention revolved around whether there should be an immediate currency union. West German Social Democrats (SPD) expressed reservations about a swift monetary and economic union and advocated for a more gradual adjustment process while maintaining

---

legal system, or language, as noted by Uhlig [2008].

5. See Juhász et al. [2023] for an insightful review.

some degree of separation between the two German regions for a certain period. By contrast, conservative leaders from the Christian Democrats (CDU), like Chancellor Helmut Kohl, advocated for a more rapid integration process. For instance, Kohl promised during the election campaign to adopt a one-to-one currency exchange rate between the East and West German Marks—a promise that contributed significantly to the conservative party's victory in the 1990 federal elections.

After the reunification, it quickly became clear that the former centrally planned East German economy was ill-equipped to compete in a globally-intergrated market economy. Evidence suggests that before 1990, only 10 percent of the former GDR's workforce were employed by firms viable at world market prices (Akerlof et al. [1991]; Collier [1991]). Moreover, East German firms struggled with low product quality and inferior brand image when entering the global market (Collier [1991]; Mertens and Müller [2022]). Exposure to high-quality West German products from world-leading manufacturers (e.g., BMW, Volkswagen, Adidas, Bosch, Braun, Siemens, Bayer, Haribo among many others) has eroded even the confidence of East German consumers in their own regional products.

The currency union between the GDR Mark and the Deutsche Mark took effect on July 1, 1990, leading to a one-to-one exchange rate between both currencies. This swift change dealt further blow to the competitiveness of East German producers. Massive wage hikes followed, causing labor costs to surge amid low productivity and rising interest rates. These gyrations resulted in an unprecedented economic breakdown in East Germany, unmatched in modern economic history. Between 1990 and 1991, East Germany lost about one third of its GDP, and non-employment rose from zero to 30 percent.[6]

Due to the severe economic breakdown, social discontent brewed quickly in the East, prompting the government to respond with substantial transfer payments.[7] Between 1991

6. Non-employment includes unemployment and employment in publicly subsidized labor market programs. See Burda and Hunt [2001] for an insightful analysis of the early years of transformation.

7. Based on national account data, we calculated that in the initial years of unification, total consumption

and 2008, the East received about 1,400 billion Euros in transfers from the West, primarily (about 900 billion) in the form of social benefits such as pensions, unemployment insurance, and active labor market policies (Paqué [2009]). The remaining 500 billion Euros mostly went to public expenditures, such as infrastructure development, repayment of GDR debt, and horizontal federal transfer schemes between the German states ("Länderfinanzausgleich"). That said, only about 100 billion Euros (approximately 2% of East Germany's cumulative nominal GDP between 1991 and 2008) were disbursed as direct investment subsidies to private firms.

Another major undertaking at the time was the overhaul of the productive capacity in the East. In a quest to revamp the Eastern production units and establish economically viable private enterprises that can operate in the integrated market economy, a publicly owned privatization agency, the *Treuhandanstalt*, founded in 1990, overtook the assets and liabilities of the state-owned enterprises of the former East Germany. The Treuhand inherited firms that employed about 4 million employees with the mandate to dismantle these enterprises, determine and privatize viable units, and shut down the unviable ones.[8] Policy makers opted for a rapid and comprehensive privatization process, with a tight deadline for the Treuhand to privatize the entire Eastern economy by the end of 1994, which the Treuhand managed to accomplish to a large extent.

Helped by these extensive efforts, the Eastern production quickly recovered in the first half of the 1990s following the massive sudden blow on impact. GDP per worker in the East increased rapidly to two thirds of the West German level by 1995 from about one third at the turn of the decade (see Figure 2.1), likely reflecting quick gains from substantial capital investment and improved reallocation of labor across new establishments. However, the convergence quickly lost steam in the second half of the decade—a striking twist that we

in East Germany exceeded East German GDP by over 70%.

8. Akcigit et al. [2023] provide an in-depth analysis of the terms imposed by the Treuhand on subsequent firm performance.

strive to comprehend through our analysis.

To summarize, the German reunification entailed a swift yet challenging process of economic integration, enormous economic dislocations in the East, and a major transformation of the Eastern economy, supported by the substantial financial transfers from the West to the East. While the early years were characterized by a rapid recovery following a large contraction on impact, the pace of economic convergence has weakened abruptly in the aftermath of this initial period of fast economic growth. It is this period of tepid convergence, which has resulted in sustained economic disparities between the East and the West, and it is the focus of our investigation.

2.3 Empirical Analysis

We use various data sets to collect key stylized facts on the German economy. We rely on aggregate data and rich firm- and product-level micro data supplied the German statistical offices. In the following, we first describe our micro data. Subsequently, we present our stylized facts that motivate our theory. Throughout our empirical analysis, we define East and West Germany without Berlin because the German data does not allow to differentiate between its former Eastern and Western parts.

2.3.1 German Firm- and Product-level Data

We use annual panel data on German manufacturing firms and their products from 1995 to 2017. The data are collected and supplied by the German statistical offices and comprise two separate statistics covering firm-level and product-level information, respectively.

Firm-level cost structure survey. We use annual firm-level panel data from the cost structure survey (KSE). The KSE is a representative and stratified survey that rotates every 4-5 years and covers a 40 percent sample of all German manufacturing firms with at least

20 employees. Firms are defined as legal units. The data contain information on firms' location, industry, employment, sales, depreciation, and several cost positions, including, among others, wage bills, R&D expenditures, and intermediate input costs by various categories. We combine this data with investment information on the population of firms with at least 20 employees from the annual investment survey to construct capital stock series using a perpetual inventory method as in Braeuer et al. (2023).[9] We clean the data from outliers by dropping the top and bottom two percent of firms with respect to the ratio of sales over capital stocks, employment, intermediate input expenditures, and wage bills. As industry classifications change multiple times over our more than two decades of data, we assign time consistent industry codes to firms using our firm-product level data as described below.

Firm-product-level data. We use annual information on sales and quantities of firms' individual products from the manufacturing sector product survey (product module). The product module covers the population of products manufactured in plants with at least 20 employees. The data links products to plants and firms. We focus our analysis on the firm-level. 90 percent of firms in our data are single-plant firms. Products are categorized by an extremely fine nine-digit product classification, defining around 6,000 distinct products in our data.[10] From the sales and quantity information, we constructed product prices. For 30 percent of products, the statistical offices do not collect quantity information. This mostly regards product categories for which quantity information is less meaningful, like the purification of products, service-like activities (e.g., printing of newspapers), or that are difficult to express and compare in terms of quantities (e.g., production of vitamins).

We clean the data from the top and bottom one percent of outliers with respect to

9. The first capital stocks are derived from the data on depreciation and aggregate information on the expected lifetime of capital goods provided by the German statistical offices.

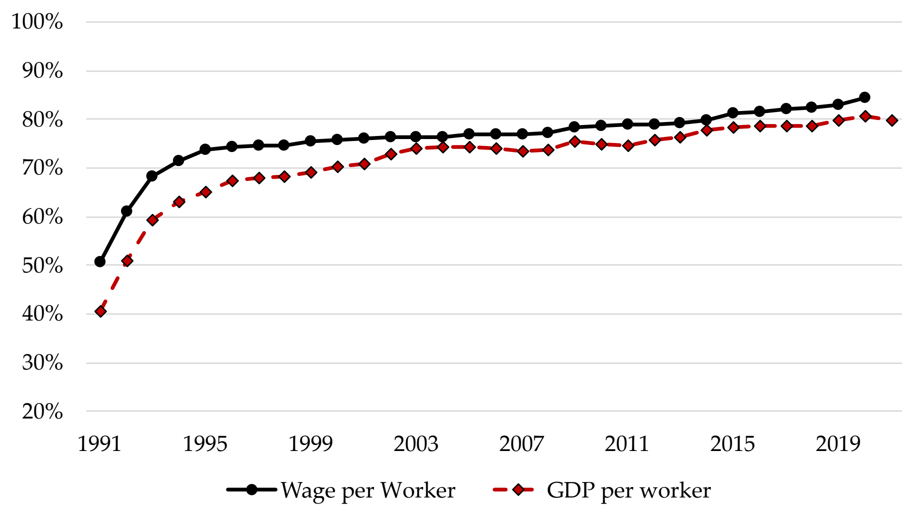
10. The product codes include an additional tenth number indicating if the product was manufactured as part of a contracted work agreement. Examples of products are "Long trousers for men, cotton", "Coats for women, chemical fibre", or "Passenger cars, petrol engine $\leq 1,000cm^3$".

product-level price deviations from the average product price. Together with the changes in firms' industry classification, product codes have been redefined twice over our period of observation. To construct a time-consistent product and industry classification, we follow Mertens and Müller [2022] and use official concordances and information on firms' product mix before and after reclassification periods to translate all product codes into the GP2002 classification scheme. The first four digits of this classification corresponds to the NACE rev. 1.1 industry definition. We therefore assign industries to firms based on firms' main production activities. Over all years, we reclassify 97 percent of products. Furthermore, for the years 2002-2007, we can compare our industry assignment with the statistical offices classification and find that we match the two-digit level in 95 percent of all cases. We drop the 3 percent of products products that we cannot reclassify. Moreover, as we study relative price differences between East- and West-Germany, we focus on products that are manufactured in both regions and by at least three firms (across both regions). This sub-sample accounts for 62 percent of production in the data.

2.3.2 Stylized facts

Fact 1: Convergence in productivity and wages was fast initially but slowed considerably in subsequent years. Figure 2.5 shows that in 1991 East German relative GDP per worker was around 40 percent of the West German level. Until the mid-1990s, the East grew to two thirds of the West German productivity level, with a commensurate rise in the relative wage level. This initial recovery is a rebound from an abysmal state of economic activity at the outset of integration, likely reflecting the quick returns from substantial capital investment, infrastructure improvements, and enhanced reallocation of labor within the East (Findeisen et al. [2021]; Heise and Porzio [2022]). Yet, subsequently, convergence tapered off. Even three decades after the fall of the Berlin wall, East German GDP per worker and the average wage level are still about 20 percent below the West German values.[11] Our goal is to study the mechanisms that underlie this sustained disparity.

Figure 2.5: GDP and gross wages per worker and hours worked in East Germany, relative to West



Notes: GDP and gross wages per worker in East Germany, relative to West Germany. Berlin is excluded. Source: national accounts of Germany - results for the German Länder.

Fact 2: After reunification, East German firms entered world markets nearly without competitive products. Based on GDR data, Akerlof et al. [1991] document that only about 10% of former GDR's workforce were employed by firms capable of competing at global market prices (e.g., Trabants vs. BMWs). Indeed, Figure 2.6 show that, between 1991 and 2020, East Germany's share of GDP (excluding Berlin) ranged from 7 to 12 percent, which remained below its population share, which hovers between 16 and 19 percent.

Analyzing the sales and quantity information derived from our detailed firm-product-level data in the manufacturing sector, a similar pattern emerges. The sales and quantity shares of East German firms, when weighted by product sales, depict only a gradual increase from 6% to 9% and from 8% to 12%, respectively, during the period from 1995 to 2017.

11. The findings also apply to measures based on hours worked, although the series are available for a shorter period (see also Figure 2.1).

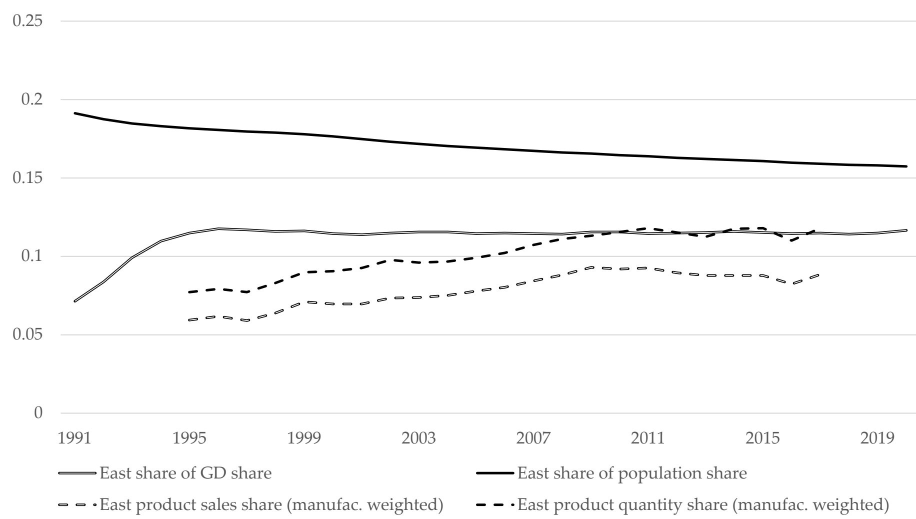


Figure 2.6: Market shares of Eastern firms

Notes: Market and population shares of East German firms. GDP shares (1991-2020) are taken from Destatis data. Population shares (1991-2020) come from the VGR der Länder. Product market shares (1995-2017) are derived from our firm-product level data. Each product-level market share is weighted with product-level sales to derive the aggregate series. Berlin is excluded.

Notably, the disparity between quantity and sales shares suggests that Eastern output prices remain below Western levels. Further analysis of price patterns is discussed below. Overall, our findings indicate a notable competitive disadvantage for the Eastern economy following reunification, with Eastern firms unable to catch up to their Western counterparts in terms of market shares. In our model, we explicitly consider the competitive disadvantage faced by Eastern firms and examine how this influenced incentives for R&D investment.

Fact 3: Within narrow nine-digit product markets, the East i) produces with lower revenue productivity, ii) produces lower priced varieties of the same product, and iii) maintains profitability by paying much lower wages. Using our firm-product level data, we run the following regression by periods:

$$y_{igt} = \beta_{East} East_{igt} + \upsilon_g + \upsilon_t + \varepsilon_{igt}. \tag{2.1}$$

$East_{it}$ is a dummy variable indicating if a firm is located in East Germany. υ_g and

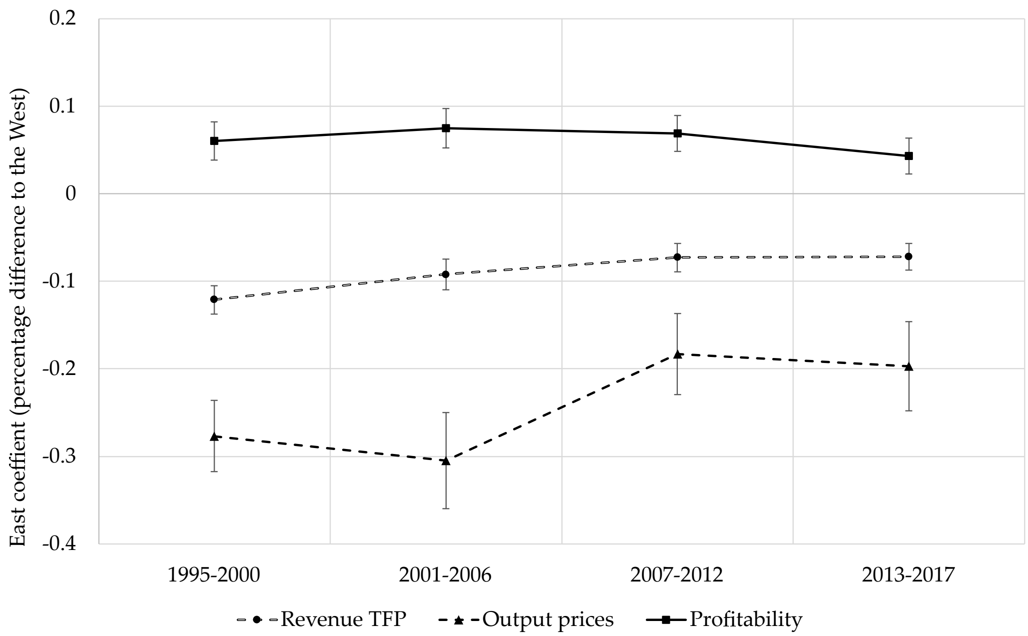


Figure 2.7: Firm-level productivity, output prices, and profitability

Notes: East-dummy coefficients from estimating eq. (2.1) by periods and using revenue total factor productivity, output prices, and profitability as dependent variables. Single-product firm sample. German manufacturing sector firm-product-level micro data. 1995-2017.

υ_t are product and year fixed effects. y_{it} indicates firm-level productivity, output prices, or profitability.[12] We define productivity as revenue total factor productivity, i.e., physical productivity times output prices. Profitability refers to sales divided by an input expenditure index. The main difference between both measures is thus that profitability uses a monetary input index, essentially measuring the return in sales per one Euro of investment. The coefficient of interest in equation (2.1) is β_{East}, which captures the average difference in y_{it} between East and West German firms, conditional on year and product fixed effects. Hence, it reflects regional firm differences of y_{it} within the same detailed nine-digit product category. We estimate equation (2.1) over four consecutive five-year periods.

Figure 2.7 summarizes the results by plotting East-dummy coefficients from estimating equation (2.1). Whereas productivity and output prices in the East are significantly lower compared to the West, firm profitability is roughly equal in both regions. This phenomenon

12. See Mertens and Müller [2022] for a more detailed analysis.

can be attributed to the considerably lower wages paid by East German firms in comparison to their West German counterparts. Eastern firms thus compete by specializing in low-wage product varieties within specific product categories, which are sold at correspondingly lower output prices. In the initial periods, we observe a convergence in productivity and output prices, which decelerates in subsequent periods. Notably, profitability remains on equal level in both regions.

The disparity between productivity and profitability is a crucial observation that aids in explaining the relatively low investment into R&D of Eastern firms. In our model, we take into account the difference between profitability and productivity by incorporating wage differentials between the two regions. By doing so, we can capture the competitive advantage enjoyed by Eastern firms due to lower wages, which reduces incentives of East German firms to invest into R&D.

Fact 4: R&D expenditures are much lower in the East, and East German R&D investment did not catch up with the West. Figure 2.8 show the R&D to GDP ratios for West and East Germany. In the West, total R&D expenditures as a percentage of GDP grew from 2.2% to 3.3%, while in the East, they increase from 1.5% to 2.3%. The differences in R&D investment become even more pronounced when examining private R&D expenditures. In the East, the ratio of private R&D expenditures to GDP only slightly increased from 0.6% to 0.9%, representing less than half of the overall R&D expenditures in East Germany. In contrast, the West German private R&D to GDP ratio rises from 1.5% to 2.3%, accounting for approximately two-thirds of West German total R&D expenditures. Figure 2.8 underscores the lack of catch-up in R&D activities by Eastern firms in comparison to their Western counterparts.

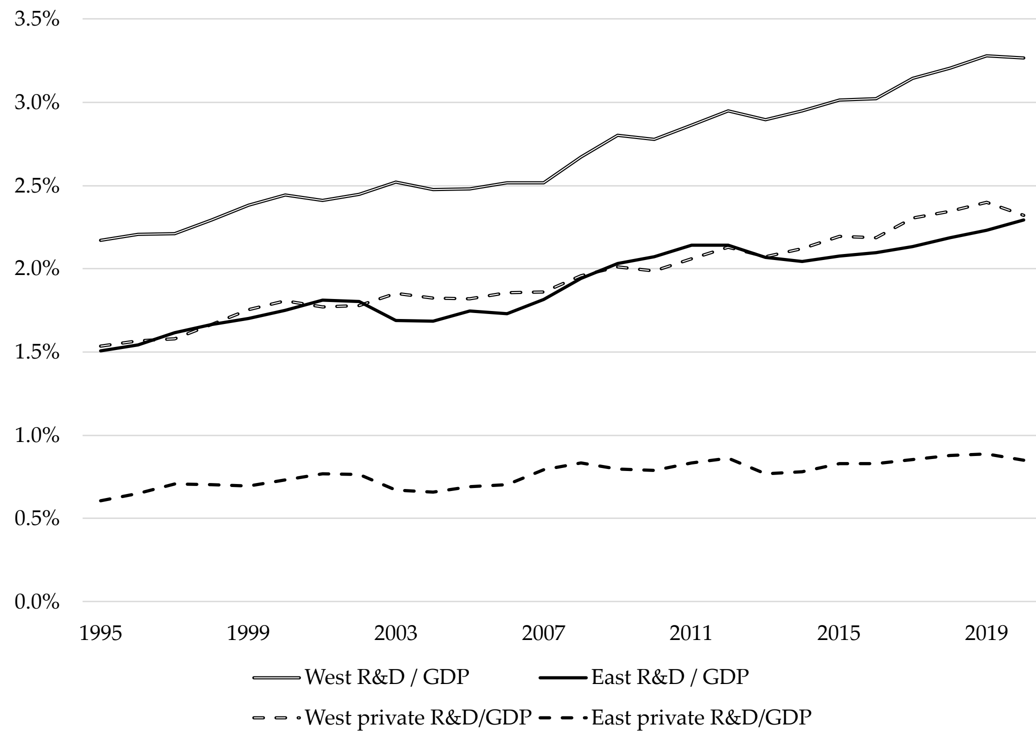


Figure 2.8: R&D expenditures in East and West Germany.

Notes: Total and private R&D expenditures over GDP for East and West Germany (1995-2020). We impute R&D expenditures for 1996 and 1998. Berlin is excluded.

2.4 Model

We build an endogenous growth model with a particular focus on competition and strategic interaction among firms. In our model, Germany is represented as a closed economy to the rest of the world. However, it is assumed that Germany consists of only two regions, West and East, which are denoted by indices $i, n \in \{w, e\}$. In the model, West and East firms compete with each other in regional product markets in terms of qualities of goods they produce. The dynamic aspect of the model stems from the fact that firms invest in R&D to improve their product quality in a step-by-step fashion as in Acemoglu and Akcigit [2012] and Akcigit and Ates [2023]. As it will be clear in next sections, dynamic strategic interactions between firms arise when firms decide for the optimal R&D investment. Furthermore, static strategic interactions also arise from the Bertrand market structure in which one West and one East firms operate, and they decide on their prices simultaneously. This structure allows us to capture realistic price dynamics and discouragement effect in innovation.

Time, denoted by $t \in [T_0, \infty)$, is continuous and runs forever. The model starts with an initial distribution of product level qualities in $T_0 = 1995$, which is calibrated to the price gap data. Representative household in each region consumes a Cobb-Douglas aggregate of tradable and nontradable goods. Firms from East and West Germany compete with each other in tradable product markets. It is assumed in the baseline model that trade is costless and frictionless.

2.4.1 Preferences

The demand system is characterized by nested constant elasticity of substitution (CES) preferences. Each region is populated by a representative household who derive utility from consumption of tradable and nontradable goods, and supply labor inelastically. Nontradable good is region specific and homogeneous, and is produced by regional firms in perfectly competitive markets. A unit measure of tradable sectors exist, which are denoted by $j \in [0, 1]$. In each sector, there are one West and one East firm who can produce a differentiated variety. These varieties are imperfect substitutes of each other with an elasticity of substitution $\sigma > 1$. In order to simplify the exposition, We assume that the elasticity of substitution between sectors j is one. This market structure delivers variable markups, which are essential to map the price gap distribution to the data.

Tradable goods are heterogeneous in quality. The quality of a good in sector j produced by region i is denoted by $q_{ij}(t)$. Quality acts like a demand shifter in household's preferences. Firms from both regions invest in R&D in order to improve their products' quality over time, as will be explained in more detail in Section 2.4.3. The demand system outlined above can be formalized by the following life-time utility maximization problem for the representative household located in region $n = w, e$

$$
\begin{aligned}
\max_{\{c^n_{ij}(t), C^{NT}_n(t), \mathcal{A}_n(t)\}} & \int_{T_0}^{\infty} e^{-\rho(t-T_0)} U_n(t) dt \\
\text{s.t. } & \dot{\mathcal{A}}_n(t) = r_n(t)\mathcal{A}_n(t) + W_n(t)L_n(t) + \Pi_n(t) + T_n(t) - E_n(t) \\
& E_n(t) \equiv \int_0^1 \Big(p_{wj}(t)c^n_{wj}(t) + p_{ej}(t)c^n_{ej}(t)\Big)\, dj + P^{NT}_n(t)C^{NT}_n(t) \\
& U_n(t) \equiv \log\left\{\left(\frac{U^T_n(t)}{\beta}\right)^{\beta}\left(\frac{U^{NT}_n(t)}{1-\beta}\right)^{1-\beta}\right\} \\
& U^T_n(t) \equiv \exp\left\{\int_0^1 \log\left(\Big(q_{wj}(t)c^n_{wj}(t)\Big)^{\frac{\sigma-1}{\sigma}} + \Big(q_{ej}(t)c^n_{ej}(t)\Big)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} dj\right\} \\
& U^{NT}_n(t) \equiv Q^{NT}_n(t)C^{NT}_n(t)
\end{aligned}
$$

In this formulation, $U^{NT}_n(t)$ denotes the utility derived from nontradable good consumption. It is the product of region specific quality of nontradable good, $Q^{NT}_n(t)$, and the amount of quantity consumed, $C^{NT}_n(t)$. In order to keep nontradable sector tractable and simple, we assume that individual qualities of tradable goods in a region spill to its nontradable sector. In particular

$$
Q^{NT}_n(t) \equiv \exp\left\{\int_0^1 \log q_{nj}(t) dj\right\} \tag{2.2}
$$

Similarly, $U^T_n(t)$ denotes the utility derived from the consumption of all tradable goods produced in the economy. Consumption of households that are located in region n for the tradable good produced in region i in sector j is denoted by $c^n_{ij}(t)$. Quality and price of this particular good is denoted by $q_{ij}(t)$, and $p_{ij}(t)$, respectively. Costless trade of goods between the regions ensures that price of a good is the same in both markets. Western and Eastern products are combined with an elasticity of substitution $\sigma > 1$ within each sector j. Finally, elasticity of substitution across sectors is assumed to be one. Final flow rate of utility is denoted by $U_n(t)$ which is a Cobb-Douglas aggregation of $U^T_n(t)$ and $U^{NT}_n(t)$, where $\beta \in (0,1)$ represents the aggregate share of tradable goods.

Households finance their total consumption expenditure, $E_n(t)$, from labor income they earn, $W_n(t)L_n(t)$, profits rebated, $\Pi_n(t)$, government transfers, $T_n(t)$, and return from financial assets they hold, $r_n(t)\mathcal{A}_n(t)$.[13] Total value of their financial portfolio is denoted by $\mathcal{A}_n(t)$. We assume that markets are incomplete in the sense that households from different regions cannot write debt contracts with each other, however, households within a region can do. We label this assumption as financial autarky. As an extension of this assumption, firms in a region are owned by the households only from the same region. Therefore, $\Pi_n(t)$ is the sum of dividends distributed by firms that produce in region n. Immediate implication of the financial autarky assumption is that trade is balanced at all times between the regions when government transfers and subsidies are absent. Moreover, the value of the financial portfolio, $\mathcal{A}_n(t)$, is equal in equilibrium to the total market value of the firms that are located in region n. Final implication of the financial autarky assumption is region specific rates of return on assets, $r_n(t)$. Financial autarky assumption allows us to abstract away from dynamic considerations and long-run trade deficit concerns. It should be noted that it is evident in the data that East experienced net trade deficit against the West, especially in the early years of unification. In the model, we capture this phenomenon by voluminous government transfers and subsidies from West to East. $T_n(t)$ is explicitly calibrated to match aggregate lump-sum transfers that is observed in the data.

Life-time utility maximization implies the standard Euler equation

$$\frac{\dot{E}_n(t)}{E_n(t)} = r_n(t) - \rho \tag{2.3}$$

We now turn to static utility maximization for a given level of expenditure $E_n(t)$ in order to derive demand functions for tradable goods.

13. Positive $T_n(t)$ stands for lump-sum government transfers to region n, whereas negative $T_n(t)$ denotes net lump-sum taxes collected from region n.

Static utility maximization Representative household in region $n = w, e$ allocates her expenditure E_n between tradable and nontradable goods by solving the following static maximization problem in each period[14]

$$U_n = \log\left\{\left(\frac{\exp\left\{\int_0^1 \log\left(\left(q_{wj}c_{wj}^n\right)^{\frac{\sigma-1}{\sigma}} + \left(q_{ej}c_{ej}^n\right)^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} dj\right\}}{\beta}\right)^{\beta}\left(\frac{Q_n^{NT}C_n^{NT}}{1-\beta}\right)^{1-\beta}\right\}$$

$$\text{s.t.} \quad \int_0^1 \left(p_{wj}c_{wj}^n + p_{ej}c_{ej}^n\right) dj + P_n^{NT}C_n^{NT} = E_n$$

Taking first order conditions and organizing terms yields the following demand functions

$$P_n^{NT}C_n^{NT} = (1-\beta)\, E_n \tag{2.4}$$

$$p_{ij}^n c_{ij}^n = \frac{\left(\frac{q_{ij}}{p_{ij}}\right)^{\sigma-1}}{\left(\frac{q_{wj}}{p_{wj}}\right)^{\sigma-1} + \left(\frac{q_{ej}}{p_{ej}}\right)^{\sigma-1}}\beta E_n \qquad \text{for } i = w, e \tag{2.5}$$

Demand function (2.5) states that representative household spends β fraction of their expenditure on each tradable sector j. Furthermore, they split this expenditure between West and East products depending on their relative quality. If the relative quality of a good q_{ij} increases, then households demand more and allocate a larger fraction of spending for that good. Finally, we can derive aggregate quantity and quality indices as follows.

$$U_n(t) = \log\left\{Q_n(t)C_n(t)\right\} \tag{2.6}$$

where $Q_n(t)$ stands for an aggregate quality index, and $C_n(t)$ denotes aggregate quantity index, which are defined below. Per-period utility flow derived from consumption of all goods

14. Time argument t is removed to save on notation in the formulation of the problem.

can simply be represented as the multiplication of aggregate quality and quantity indices. Aggregate quality index can be defined as

$$Q_n(t) \equiv \left(Q^T(t)\right)^{\beta} \left(Q_n^{NT}(t)\right)^{1-\beta}$$
$$Q^T(t) \equiv \exp\left\{\int_0^1 \log q_j(t) dj\right\}$$
$$q_j(t) \equiv \left[q_{wj}(t)^{\frac{\sigma-1}{\sigma}} + q_{ej}(t)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

where $q_j(t)$ is the aggregated quality of West and East goods in tradable sector j, and $Q^T(t)$ corresponds to the aggregate quality of all tradable goods. $Q_n^{NT}(t)$ is the quality of nontradable goods in region n which is given by (2.2). Corresponding to these quality indices, we have the following set of definitions for quantity indices

$$C_n(t) \equiv \left(\frac{C_n^T(t)}{\beta}\right)^{\beta} \left(\frac{C_n^{NT}(t)}{1-\beta}\right)^{1-\beta}$$
$$C_n^T(t) \equiv \exp\left\{\int_0^1 \log c_j^n(t) dj\right\}$$
$$c_j^n(t) \equiv \left[\left(\tilde{q}_{wj}(t) c_{wj}^n(t)\right)^{\frac{\sigma-1}{\sigma}} + \left(\tilde{q}_{ej}(t) c_{ej}^n(t)\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$
$$\tilde{q}_{ij}(t) \equiv \frac{q_{ij}(t)}{q_j(t)}$$

where $\tilde{q}_{ij}(t)$ represents the quality of tradable good i in sector j relative to the quality index of its sector j. These indices will prove useful when calculating welfare of households. It can be seen from (2.6) that the utility of representative household increases in the quality of goods consumed. This is the main source of the increase in welfare of households in the model economy. Finally, aggregate price index $P_n(t)$ corresponding to the quantity index $C_n(t)$ can be defined as

$$P_n(t) \equiv \frac{E_n(t)}{C_n(t)}$$

This aggregate price index equals to the following expression

$$
\begin{aligned}
P_n(t) &\equiv \left(P^T(t)\right)^{\beta} \left(P_n^{NT}(t)\right)^{1-\beta} \\
P^T(t) &\equiv \exp\left\{\int_0^1 \log p_j(t) dj\right\} \\
p_j(t) &\equiv \left[\left(\frac{p_{wj}(t)}{\tilde{q}_{wj}(t)}\right)^{1-\sigma} + \left(\frac{p_{ej}(t)}{\tilde{q}_{ej}(t)}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}
\end{aligned}
$$

Time varying exogenous variables or parameters We solve the model on the transition path equilibrium in order to capture convergence dynamics. Some parameters and exogenous variables are time varying in the model. A certain functional form is assumed for such variables which are parameterized by only three additional parameters. Any exogenous time varying variable $x_i(t)$ follows the following time path

$$
x_i(t) = x_i^{\star} + \left(x_{i,0} - x_i^{\star}\right) e^{-\nu_i^x (t-T_0)} \tag{2.7}
$$

This process is parameterized by only three unknowns. Firstly, $x_{i,0}$ denotes the initial value of the variable at $t = T_0$. Secondly, $x_i^{\star}$ is the parameter to which the variable $x_i(t)$ converges over time. Finally, $\nu_i^x > 0$ is the parameter that governs the speed of convergence of $x_i(t)$ to $x_i^{\star}$. In particular, the half-life of $x_i(t)$ can be found as $\frac{\log 2}{\nu_i^x}$. The condition that ν_i^x being positive ensures that $x_i(t) \to x_i^{\star}$ as $t \to \infty$ from any initial condition of $x_i(T_0) = x_{i,0}$. Which exogenous variables or parameters follow this particular form are explained in later sections.

Regional labor supply In order to capture migration dynamics that took place between the regions after the unification, we assume that regional labor supplies are possibly time varying, and exogenous. That is, we abstract away from endogenous migration decisions of agents between the regions. In endogenous growth models, scale effects are important in the sense that when labor is used for R&D purposes, the growth dynamics of the economy might be responsive to the amount of labor supply in the economy (Jones [1995]). Therefore, instead of explicitly assuming constant labor supplies in West and East, we assume that $L_w(t)$ and $L_e(t)$ follow the process described by equation (2.7). Furthermore, total labor supply in Germany is normalized to one at all times, i.e.

$$L_w(t) + L_e(t) = 1$$

This assumption leaves five parameters that govern the evolution of regional labor supplies. The first two of these are $L_{w,0}$ and $L_{e,0}$ such that $L_{w,0} + L_{e,0} = 1$, which stands for the initial level of labor supply in regions at time T_0. Other two parameters are their limit value, $L_w^\star$ and $L_e^\star$ such that $L_w^\star + L_e^\star = 1$. The last parameter is ν^L determining the shape of the evolution over time. The reason as to why convergence speed parameter does not change across the regions is because the sum of labor supplies is normalized to one at all times.[15] Calibration of these five parameters is performed by fitting the equation (2.7) to the aggregate employment time series data. The details are explained in Section 2.5.1 in which the calibration strategy is discussed.

2.4.2 Production

Labor is the only factor of production in the model. As will be explained in more detail below, production technologies for nontradable and tradable goods are linear in labor. All firms in a region are assumed to have access to an aggregate and region specific production technology, which might be heterogeneous across the regions. There are two reasons for such a difference. First of all, the model abstracts away from capital accumulation within regions and capital transfers across regions. However, as the model is designed to reflect

15. Substituting equation (2.7) into the normalization equation implies that for the total labor supply to stay constant over time, it has to be either both $\nu_w^L = 0$ and $\nu_e^L = 0$, or $\nu_w^L = \nu_e^L \neq 0$. This common speed parameter which is different than zero is denoted by ν^L.

the changes on the convergence process of East to West after the unification, such short term considerations become important for the model to capture realistic dynamics that are observed in the data. It is also natural to assume that human capital in the East was different than that in the West shortly after the unification took place. Although the initial human capital differences were possibly huge in the short run, they could have converged to similar levels over time. Because of all these reasons, we model the aggregate productivity of a region as an exogenous time-varying process which is also flexible enough to capture the convergence dynamics observed in the data.

Productivity spillovers First kind of spillovers between the regions stems from the possibility that labor in the East can learn latest production techniques from the West. This assumption can be justified by large gross migration flows between the regions after the unification. As soon as the borders between the GDR and West Germany was removed, there had been voluminous interactions between the members of the two regions. In the model, these interactions are captured as exogenous changes in aggregate productivity level of East German labor force over time.

Aggregate labor productivity in region i at time t is denoted by $A_i(t)$. First of all, we normalize the West aggregate productivity to be one in all periods, i.e. $A_w(t) = 1$ for all $t \geq T_0$. Then, it is assumed that aggregate East productivity $A_e(t)$ follows the exogenous process given by (2.7). Note that $A_e(t)$ can also be interpreted as the aggregate efficiency of East German labor relative to their Western counterparts. Finally, we assume that East productivity converges to Western level in the long run, i.e. $A_e^\star = 1$. Therefore, two parameters, $A_{e,0}$ and ν_e^A, determine the time evolution of $A_e(t)$ completely.[16]

16. Under these assumptions, $A_e(t)$ is given by

$$A_e(t) = 1 + (A_{e,0} - 1)\, e^{-\nu_e^A (t - T_0)}$$

Nontradable goods A homogeneous nontradable good specific to a region $i = w, e$ is produced by perfectly competitive firms in the region. Production technology is linear in labor, $Y_i^{NT}(t) = A_i(t)L^{NT}(t)$. Firms solve the following static problem at each instant

$$\max_L P_i^{NT}(t)A_i(t)L - W_i(t)L$$

by taking prices $P_i^{NT}(t)$ and wages $W_i(t)$ as given. Given that representative household spends $1-\beta$ fraction of their total expenditure on nontradable goods, we can show that labor allocated for nontradable good production is given by $L_i^{NT}(t) = (1-\beta)\, E_i(t)/W_i(t)$.

Tradable goods In every sector $j \in [0, 1]$, West and East firms engage in Bertrand price competition. Production function is given by

$$y_{ij}(t) = A_i(t)L_{ij}^T(t)$$

where $L_{ij}^T(t)$ denotes labor allocated in region i for the production of sector j good. Profit maximization problem of a firm in any sector j is given by

$$\max_p \left[p - \frac{W_i(t)}{A_i(t)}\right] y\,(p)$$
$$\text{s.t. } y(p) = \frac{1}{p}\frac{\left(\frac{q(t)}{p}\right)^{\sigma-1}}{\left(\frac{q(t)}{p}\right)^{\sigma-1} + \left(\frac{q_{-1}(t)}{p_{-1}(t)}\right)^{\sigma-1}}\beta E(t)$$

where $E(t)$ equals total expenditures, $E(t) = E_w(t) + E_e(t)$, $q(t)$ denotes the quality of the focal firm, $q_{-1}(t)$ and $p_{-1}(t)$ denote quality and price of firm's competitor, respectively. The demand function can be derived by summing region specific demands given by equation (2.5). Solution to this maximization problem gives the optimal pricing rule as a markup over the marginal cost as usual. However, this markup is not constant, and in particular it increases in the market share of the firm. Denoting by $s_{ij}(t) \in [0, 1]$ the market revenue share of the firm, we can show that prices and markups are equal to

$$p_{ij}(t) = \mu\left(s_{ij}(t)\right)\frac{W_i(t)}{A_i(t)} \tag{2.8}$$

$$\mu\left(s\right) = \frac{\varepsilon(s)}{\varepsilon(s)-1} \tag{2.9}$$

$$\varepsilon(s) = s + (1-s)\,\sigma \tag{2.10}$$

Residual demand elasticity for firm i is a function of the size of the firm measured by its market share $s_{ij}(t)$. As the market share of the firm rises, the demand that the firm faces gets more inelastic, which allows the firm to charge higher markups. Indeed, the price elasticity of demand is a weighted average of across-sector elasticity one and within-sector elasticity $\sigma > 1$, weights being the market share of the firm. As s increases, this average declines and markup of the firm increases.

Final equation that determines the market shares in equilibrium as a function of relative qualities can be derived after replacing optimal price (2.8) into demand (2.5), and using the fact that market shares of two firms in a sector sum up to one. Market share of firm i when its competitor is denoted by $-i$ can be solved from the nonlinear equation below

$$\frac{s_{ij}(t)}{1-s_{ij}(t)} = \left(\frac{q_{ij}(t)}{q_{-ij}(t)}\right)^{\sigma-1}\left(\frac{\mu\left(s_{ij}(t)\right)W_i(t)/A_i(t)}{\mu\left(1-s_{ij}(t)\right)W_{-i}(t)/A_{-i}(t)}\right)^{1-\sigma} \tag{2.11}$$

As is apparent from equation (2.11), a firm's market share increases as it obtains quality advantage over its competitor. Since markup increases with market share, higher quality firms can charge higher markups and higher prices for their goods. This theoretical positive relationship between good quality and prices is exploited to discipline the distribution of quality gaps between East and West firms, which is otherwise unobserved, by matching the realized price gap distributions with the data. Finally, profits and labor demands can be derived as

$$
\pi_{ij}(t) = \left[1 - \mu\left(s_{ij}(t)\right)^{-1}\right] s_{ij}(t)\beta E(t)
$$

$$
L^T_{ij}(t) = \frac{\frac{s_{ij}(t)}{\mu(s_{ij}(t))}\beta E(t)}{W_i(t)}
$$

2.4.3 Innovation and knowledge spillovers

For simplicity, the model is abstracted away from firm entry as opposed to standard step-by-step innovation models. Incumbent firms compete with each other only at the quality margin. They increase quality of their products by investing in R&D. Per successful innovation, it is assumed that the quality of product jumps from its level of q to λq, where $\lambda > 1$ is called innovation step size. This structure allows us to define quality gaps which turns out to be the only state variable of a given sector. Prices, markups, labor demands, market shares are functions of the quality gap between the competitors.

The quality gap in terms of the number of steps between two firms in a sector is denoted by $m \in \{-\bar{m}, \ldots, -1, 0, 1, \ldots, \bar{m}\}$. Since step size is constant, the integer m gives a measure of relative qualities within a sector

$$
m_{ij}(t) = \log_\lambda \frac{q_{ij}(t)}{q_{-ij}(t)} \tag{2.12}
$$

We restrict the domain of m with a high enough constant integer $\bar{m}$. Thus, $\bar{m}$ gives the highest possible quality gap that any firm can achieve against its competitor. At this boundary, firms never invest in R&D anymore. This restriction allows us to solve the equilibrium numerically. Replacing m given by definition (2.12) into the market share equation (2.11) and markup equations (2.9) and (2.10) allows market shares and markups to be expressed as functions of only the quality gap m

$$\frac{s_i(m,t)}{1-s_i(m,t)} = \lambda^{(\sigma-1)m}\left(\frac{\mu\left(s_i(m,t)\right)W_i(t)/A_i(t)}{\mu\left(1-s_i(m,t)\right)W_{-i}(t)/A_{-i}(t)}\right)^{1-\sigma} \tag{2.13}$$

R&D technology available to firms is specified as follows. A firm in region i that chooses a rate of innovation of x has to employ $\alpha_i \frac{x^\gamma}{\gamma}$ many researchers. x is the Poisson arrival rate of a successful innovation which improves the quality of the firm with one step. Importantly, it is assumed in the model that R&D technology might differ across regions. Several possible explanations can be proposed for such an assumption. First of all, Western economy was built on a market economy structure whose main driver of growth is innovation. However, Eastern economy lacked such incentives for R&D before the unification. Secondly, institutions differ across regions, as it is difficult to argue the presence of pro-innovation institutions in the East. The main reason is that allocation of resources had been commanded by central government in the East. However, in the West, entrepreneurs have been profit maximizers, and it is natural to assume that institutions in the West had emerged for the needs of an innovation-led market economy.

Quality spillovers We assume that follower firms in terms of quality may receive a quality spillover from the market leader with a Poisson rate of $\delta \geq 0$. Quality spillovers are the second type of spillovers present in the model. Upon arrival of the spillover, follower firm draws an integer for how many steps it catches up with the leader from a Pareto distribution with a shape parameter $\theta_{\text{shape}} = 0.01$ and location parameter $\theta_{\text{location}} = 1$. That is, the minimum possible catch up is only one step, whereas the maximum quality spillover is restricted such that the follower firm can jump at most to the same quality level of the market leader, but can never leapfrog it.

Firm values Let $V_i(m,t)$ denote the discounted flow of future cash flows of a firm located in region i with $m \in \{-\bar{m}, \dots, \bar{m}\}$ quality steps ahead of its competitor. Hamilton-Jacobi-Bellman (HJB) equation can be written as follows

$$\begin{aligned}
r_i(t)V_i(m,t) &= \left[1 - \mu\left(s_i(m,t)\right)^{-1}\right] s_i(m,t)\beta E(t) \\
&+ \max_x \left\{ -\left(1-\zeta_i(t)\right) W_i(t)\alpha_i \frac{x^\gamma}{\gamma} + x\left[V_i(m+1,t) - V_i(m,t)\right] \right\} \\
&+ x_{-i}(-m,t)\left[V_i(m-1,t) - V_i(m,t)\right] \\
&+ \mathbf{1}_{\{m \neq 0\}}\delta \left[\sum_{n=1}^{|m|} f\left(|m|, n\right) \left[\mathbf{1}_{\{m>0\}} V_i(m-n,t) + \mathbf{1}_{\{m<0\}} V_i(m+n,t)\right] - V_i(m,t) \right] \\
&+ \frac{\partial V_i(m,t)}{\partial t}
\end{aligned} \tag{2.14}$$

The left hand side of (2.14) represents the safe return which must be equal to the return if the firm is operated. The first line of the right hand side is the flow rate of profits. The second line is the maximization for optimal innovation rate which is to be solved by the firm. It equals to the expected return from R&D, which is the increase in value from increasing the gap from m to $m+1$ minus the cost of R&D. Region specific R&D subsidy rate is given by $\zeta_i(t) \in [0,1]$. Because the input of R&D is labor, total employment needed to generate an innovation rate of x is multiplied by the wage rate in the region. Third term is the expected loss in case of a successful innovation by the competitor. In that case the gap of the firm declines from m to $m-1$. The forth term is the expected change in firm value in case of a knowledge spillover. The size of the spillover is assumed to depend on the position of the firm on the quality ladder. In particular, $f\left(|m|, n\right)$ gives the probability of jumping by n steps if the absolute gap between the firms is $|m|$. If the firm is laggard, then spillover increases firm value. On the other hand, if the firm is the leader in the market, then spillover decreases firm valuation. Finally, the last term represents the change in firm value due to changes in aggregate variables. First order condition to firm's maximization problem gives the innovation rate chosen by the firm

$$x_i(m,t) = \left[\frac{V_i(m+1,t) - V_i(m,t)}{\left(1 - \zeta_i^R(t)\right) \alpha_i W_i(t)} \right]^{\frac{1}{\gamma-1}} \tag{2.15}$$

Strategic interactions between the firms in a sector arise because of the step-by-step structure of innovations. Although imperfect substitution between firms is assumed, firms that are very close to each other on the quality ladder invests the most in R&D. The reason stems from the steepness of the profit function. If σ is high, once a firm captures quality leadership in the market, firm profits increase by a large margin. This huge increase in profits incentivizes firms to invest more in R&D. On the other hand, follower firms that are far away from their competitors on the quality ladder have to invest much more to catch up with the market leader and make significant profits. This increases the cumulative cost of R&D for these firms, having them invest less in R&D. Due to rational expectations and the fact that there are only two firms in the market, leader firm expects that the laggard firm is discouraged from R&D. As a result, the competitive pressure on leader firm decreases, and it also starts investing less in innovation.

2.4.4 Quality gap distribution

The quality gap distribution is the aggregate state of the model that moves endogenously over time. Its evolution in equilibrium is determined by the rates of innovations chosen by the firms, and quality spillovers. As we solve the model for transition periods as well, initial quality gap distribution has to be given explicitly. We estimate this distribution from the price gap distribution in 1995 observed from the data.

Let $\psi(m,t)$ denote the mass of sectors in which Western firms have a lead of $m \in \{-\bar{m}, \ldots, \bar{m}\}$ steps. The mass of sectors in which East firms have a gap of m equals to $\psi(-m,t)$. Kolmogorov forwards equations for $m = -\bar{m}+1, \ldots, \bar{m}-1$ can be derived using the inflows and outflows due to innovation choices as follows

$$
\begin{aligned}
\dot{\psi}(m,t) =& \psi(m-1,t)x_w(m-1,t) + \psi(m+1,t)x_e\left(-(m+1),t\right) \\
&+ \delta\left[\sum_{n=1}^{\bar{m}-|m|} \mathbf{1}_{\{m\geq 0\}} f(m+n,n)\psi(m+n,t) + \mathbf{1}_{\{m\leq 0\}} f\left(|m-n|,n\right)\psi(m-n,t)\right] \\
&- \psi(m,t)\left[x_w(m,t) + x_e(-m,t) + \mathbf{1}_{\{m\neq 0\}}\delta\right]
\end{aligned}
\tag{2.16}
$$

In words, the change in the mass of sectors with Western gap m equals to inflows minus outflows. In particular, first line gives the inflow due to West and East firm innovations. Second line represents the inflow due to knowledge spillovers depending on whether m is negative or positive. Finally, the last line denotes total outflow because of innovations in the sectors with quality gap m and knowledge spillovers.

Boundaries $\bar{m}$ and $-\bar{m}$ need special treatment as it is assumed by construction that rate of innovation is zero if a firm has a quality gap of $\bar{m}$ against its competitor. The reason is that leader firms choose a zero rate of innovation at the boundaries because it is the maximum gap they can achieve. Kolmogorov forward equations for them can be derived as follows

$$
\dot{\psi}(\bar{m},t) = \psi(\bar{m}-1,t)x_w(\bar{m}-1,t) - \psi(\bar{m},t)\left[x_e\left(\bar{m},t\right) + \delta\right] \tag{2.17}
$$

$$
\dot{\psi}(-\bar{m},t) = \psi(-\bar{m}+1,t)x_e(\bar{m}-1,t) - \psi(-\bar{m},t)\left[x_w(-\bar{m},t) + \delta\right] \tag{2.18}
$$

Finally, the total mass of sectors is normalized to one

$$
\sum_{m=-\bar{m}}^{\bar{m}} \psi(m,t) = 1, \quad \forall t
$$

2.4.5 Market clearing conditions

For notational simplicity, we define region specific quality gap distributions $\psi_w(m,t) \equiv \psi(m,t)$ and $\psi_e(m,t) \equiv \psi(-m,t)$, where $\psi_i(m,t)$ denotes the mass of sectors in which firm from region i has a quality gap of m.

Labor and asset markets in each region are cleared in every period. Total labor demanded for R&D investment in region i can be found as

$$L_i^R(t) = \sum_{m=-\bar{m}}^{\bar{m}} \psi_i(m,t)\alpha_i \frac{x_i(m,t)^\gamma}{\gamma} \tag{2.19}$$

Remaining labor $L_i(t) - L_i^R(t)$ is used for production purposes. Using labor demands for tradable and nontradable goods production in a region, we can derive labor market clearing condition in region $i = w, e$ as follows

$$W_i(t)\left(L_i(t) - L_i^R(t)\right) = \underbrace{\left(\sum_{m=-\bar{m}}^{\bar{m}} \psi_i(m,t)\frac{s_i(m,t)}{\mu\left(s_i(m,t)\right)}\right)\beta E(t)}_{\substack{\text{Labor demand for tradable good} \\ \text{production}}} + \underbrace{(1-\beta)\, E_i(t)}_{\substack{\text{Labor demand for} \\ \text{nontradable good production}}} \tag{2.20}$$

Asset market clearing condition combined with the assumption of financial autarky implies that total expenditures of a region on consumption goods equals to total income. That is,

$$E_i(t) = W_i(t)L_i(t) + \Pi_i(t) + T_i(t) \tag{2.21}$$

Asset market clearing condition (2.21) can be further utilized to derive expressions for total expenditures, $E_i(t)$, and GDP of a region, $\mathcal{Y}_i(t)$, in terms of average market shares of region's firms and net transfers to the region. Total GDP of a region, $\mathcal{Y}_i(t)$, is defined as the sum of total value added generated by both tradable and nontradable firms

$$\mathcal{Y}_i(t) \equiv (1-\beta)\, E_i(t) + \sum_{n\in\{w,e\}} \left(\sum_{m=-\bar{m}}^{\bar{m}} s_i(m,t)\psi_i(m,t)\right)\beta E_n(t)$$

Total net transfers to a region is defined as the sum of total R&D subsidies given to region's firms and lumpsum transfers/taxes, $\mathcal{T}_i(t) \equiv \zeta_i^R(t)W_i(t)L_i^R(t) + T_i(t)$. Given these definitions, and defining the average market share of firms as $\mathcal{S}_i(t) \equiv \sum_{m=-\bar{m}}^{\bar{m}} s_i(m,t)\psi_i(m,t)$, we can derive the following equations

$$E_i(t) = \mathcal{S}_i(t) + \frac{1}{\beta}\mathcal{T}_i(t) \tag{2.22}$$

$$\mathcal{Y}_i(t) = \mathcal{S}_i(t) + \frac{1-\beta}{\beta}\mathcal{T}_i(t) \tag{2.23}$$

Government budget constraint It is assumed that government does not borrow to finance its expenditures. Thus, sum of net transfers to regions equals to zero, $\mathcal{T}_w(t)+\mathcal{T}_e(t) = 0$.

2.5 Calibration

In this section, We present the details of the model calibration to the transitional dynamics of economic convergence of East since the past 30 years. We also discuss key mechanisms that shape the dynamics of this convergence through the lens of the model.

2.5.1 Calibration

Our calibration strategy involves several steps. Some model parameters are assumed to be time varying, and need to be externally calibrated. In order to solve the model for transitional dynamics, initial quality gap distribution in time $T_0 = 1995$ must also be calibrated. The model suggests a very close relationship between product level prices and qualities. We exploit the full price gap distribution between East and West firms in the first year of the data, which is 1995, in order to back out initial quality gap distribution. Given initial quality gap distribution and other externally calibrated parameters, we finally calibrate the remaining parameters jointly by targeting certain moments that we obtain from the data with the moments generated by the model.

Parameter	Description	Value
ρ	Time discount rate	0.025
γ	R&D curvature	2
β	Tradable sector expenditure share	0.25
θ_{shape}	Shape parameter for Pareto distribution of knowledge spillovers	0.01
θ_{location}	Location parameter for Pareto distribution of knowledge spillovers	1
$L_{wt} + L_{et}$	Total labor of Germany, $\forall t$	1
$L_{w,0}$	West labor in 1995	0.832
$L_{e,0}$	East labor in 1995	0.168
$L_w^\star$	West labor in BGP	0.867
$L_e^\star$	East labor in BGP	0.133
$\log 2/\nu^L$	Half-life of labor	10.529 years
$T_{e,0}$	Lumpsum transfer to East in 1995	0.022
$T_e^\star$	Lump-sum transfer to East in BGP	0
$\log 2/\nu_e^T$	Half-life of lump-sum transfer to East	10 years
$\zeta_w^\star$	West R&D subsidy rate in BGP	0
$\zeta_e^\star$	East R&D subsidy rate in BGP	0

Table 2.1: Externally calibrated parameters

External calibration. The unit of time in the model corresponds to a year. Table 2.1 lists externally calibrated parameters. Annual time discount rate is set to 2.5%. The share of tradable goods in final consumption, β, is set to 0.25 as observed in the data. The curvature parameter of the R&D cost function γ, which governs the elasticity of innovative output to R&D, is set to 2 following the conventional estimates in the literature. (Acemoglu et al. [2018])

The total labor force in German economy is normalized to one, distributed at a ratio of about one to five across East and West initially in 1995. To account for migration flows and the associated changes in the size of regional labor force observed over the period of analysis, we fit the exogenous path (2.7) discussed in Section 2.4.1 to the observed path of changes in regional labor forces in the data while maintaining the aggregate size constant at one at all times, i.e. $L_W(t) + L_e(t) = 1$. The labor force in West Germany increased from 83% of total Germany employment in year 1995 to 86% in year 2015. On the flip side of the coin, East labor force declined from 17% of total employment in Germany in 1995 to 14% in 2015. It can be argued that there was net migration in terms of fractions from East to West for the first 20 years of unification. In order to capture these dynamics, we estimate five parameters that govern the exogenous evolution of regional labor forces by nonlinear least squares. Estimated values for labor parameters are as follows, $L_{w,0} = 0.832, L_w^\star = 0.867, L_{e,0} = 0.168, L_e^\star = 0.133, \nu^L = 0.066$. The half-life of convergence in labor series is $\log 2/\nu^L = 10.5$ years.

Benchmark calibration exercise takes into account the sizable support provided by the West to the Eastern consumers in the form of lump-sum consumption subsidy. In the data, this subsidy amounts to about 600 billion Euros over the period 1991-2009, though we lack information on its decomposition over time. This consumption subsidy corresponds to $T_e(t)$ variable in the model. Similar to regional labor force evolution over time, we assume that $T_e(t)$ follows the same path given by (2.7) with its own parameters to be calibrated, i.e. $T_{e,0}, T_e^\star, \nu_e^T$. As we have only a single data point, which is the total amount of subsidies between 1991-2009, we externally set two of three parameters. Particularly, we assume that these consumption subsidies do not continue in the limit, hence $T_e^\star = 0$ is taken. For the rate of convergence, we assume that the half life of subsidies is around 10 years. Corresponding value of the convergence parameter is found to be $\nu_e^T = \frac{\log 2}{10} = 0.069$. The remaining parameter $T_{e,0}$, which is the value of lump-sum subsidy in year 1995 is solved so that the total amount of subsidy equals to 600 billion Euros between the years 1991-2009. This exercise results in $T_{e,0} = 0.022$.

For the time evolution of regional R&D subsidy rates, we also assume it follows the smooth path given by (2.7) with parameters $\zeta_{i,0}$ as the initial value in year 1995, $\zeta_i^\star$ as the value in the limit (BGP), and ν_i^R as the speed of convergence, for each region $i = w, e$. we externally set the R&D subsidy rate in the BGP equilibrium to zero, i.e. $\zeta_w^\star = \zeta_e^\star = 0$. Remaining parameters that govern the evolution of R&D subsidy rates in the transition is jointly calibrated with other parameters as discussed below.

Internal calibration and identification. There are 11 parameters and initial quality gap distribution to be jointly estimated. Initial quality gap distribution is a function which maps quality gap m to the mass of sectors. We set $\bar{m} = 100$ to make sure that quality gap distribution never accumulates at the boundaries both in transition and in BGP equilibrium. Therefore, initial quality gap distribution consists of $2\bar{m}+1 = 201$ unknowns whose sum is one.

Initial quality gap distribution is identified from the observed price gap distribution between East and West in 1995. The link between quality gap distribution and price gap distribution is established by the optimal pricing rule of individual firms given by Equation (2.8), and the market share equation (2.11). Define the log gap $\hat{x}_{ej}(t) = \frac{x_{ej}(t)}{x_{wj}(t)}$ for any generic variable $x_{ij}(t)$ where j denotes a 9-digit sector. Then taking logs of Equations (2.8) and (2.11) and arranging yields the following relationships

$$\hat{p}_{ej}(t) = \hat{\mu}\left(\hat{s}_{ej}(t)\right) + \hat{W}_e(t) - \hat{A}_e(t) \tag{2.24}$$

$$\hat{s}_{ej}(t) = (\sigma - 1)\left(\hat{q}_{ej}(t) - \hat{p}_{ej}(t)\right) \tag{2.25}$$

Equations (2.24) and (2.25) constitute our baseline argument for identification. For the identification of initial quality gap distribution, we need a mapping from observed $\hat{p}_{ej}(1995)$ to unobserved $\hat{q}_{ej}(1995)$ for each of 9-digit sectors. This relationship has to be established with the whole set of estimated parameters jointly, as the relationship between prices and quality gaps depend on regional wage gaps $\hat{W}_e(t)$ and regional productivity gap $\hat{A}_e(t)$. Therefore, matching relative wages in 1995 is necessary for the identification of initial quality gap distribution in year 1995. Furthermore, relative aggregate productivity is not directly observed in the data. However, per capita income of East relative to West is strongly related to aggregate productivity of East relative to West. Therefore, we also target relative per capita income of East households in year 1995 in our calibration exercise. Having these two quantities, $\hat{W}_e(t)$ and $\hat{A}_e(t)$, quality gap $\hat{q}_{ej}(1995)$ can be directly solved from observed price gaps $\hat{p}_{ej}(1995)$ for all the sectors j. Given λ, we can finally map log quality gaps to integers m by rounding them to the closest integer. A very large choice of $\bar{m}$ such as $\bar{m} = 100$ makes sure that none of the sectors accumulate around the boundaries. In fact, most of sectors in 1995 are estimated to lie between $m = -10$ and $m = 10$ in 1995.

Above discussion also hints about the main identification argument of the paper except the identification of initial quality gap distribution. Close examination of equations (2.24) and (2.25) reveals that as a region's economy grows over time, we might observe opposing effects on firm level prices. For the rest of the argument, we focus on convergence of East to West. One reason as to why East converge to Western levels in terms of per capita income of East households is improvements in regional productivity of East, $A_e(t)$, over time. The effect of increased relative productivity in East on firm level prices is negative, as higher productivity implies lower marginal cost. On the other hand, another reason as to why East converges to West is improvements in relative qualities of Eastern goods. Higher relative quality of East goods help them capture higher market shares in tradable sector against their Western competitors, which translates into higher income and purchasing power for Easterners. As can be seen from equation (2.25), higher relative quality increases firm's market share in its sector. Higher market share in turn increases firm's markups and hence its relative prices. Therefore, we can conclude that the identification of mechanisms responsible for the observed convergence of East, albeit slow, stems from the fact that productivity-induced and quality-induced convergence have opposing effects on firm level prices and price gap distribution overall. We should also note secondary effects of these changes. Another determinant of firm level prices is wages. Therefore, it is imperative for our calibration strategy to target a moment from price gap distribution in a future year such as 2015. This moment is the average log price gap between East and West firms in 2015.

Identification of R&D cost parameters α_w and α_e comes from the direct effect of these variables on total R&D spending in regions as a share of total German economy. In these models, the indirect effect usually dominates. That is, higher α corresponds to costlier R&D for firms, hence lower spending on R&D as a fraction of their revenue, despite the fact that R&D is costlier. Therefore, α and total regional R&D spending as a share of GDP moves in opposite directions. In order to capture this identification argument, we target average West and East R&D spending as a share of total German GDP between the years 1995 and 2015.

R&D subsidy parameters, $\zeta_{w,0}, \zeta_{e,0}, \nu_w^\zeta$, and ν_e^ζ are identified from matching four data points, total R&D subsidies to West and East as a fraction of German GDP in years 1995 and 2015. Specifically, two data points that correspond to initial year 1995 allows us to pin down $\zeta_{w,0}$ and $\zeta_{e,0}$, and future data points in year 2015 helps the identification of the time evolution of the R&D subsidy parameters, which are governed by ν_w^ζ and ν_e^ζ.

Remaining target moments are East relative GDP per capita in 1995 and 2015, and East relative wage in 1995 and 2015. First of all, these four data points helps the calibration of the evolution of aggregate productivity of Eastern laborers, $A_e(t)$. Its BGP value in the long run is normalized to the Western level of one as explained previously. The remaining two parameters, $A_{e,0}$ and ν_e^A, are closely associated with East relative GDP per capita in years 1995 and 2015. Finally, East relative wage in years 1995 and 2015 are helpful in identifying the remaining two parameters, λ the innovation step size, and σ the elasticity of substitution. It should be noted that σ, and relative markups and wages are closely related. Higher elasticity of substitution allows the more quality advanced region, in this case it is the West, to charge higher markups than the East. This is because, as the elasticity of substitution gets higher, market leaders capture larger market share than before, allowing them to charge higher markups. This dampens Western wages more than wages in the East. The innovation step size is also associated with relative wage difference between the regions conditional on quality gaps in terms of steps m. For any level of m, higher λ implies larger difference in qualities, which in turn translates into larger difference in market shares and markups for leader firms.

Let Θ, denote the set of jointly calibrated parameters

$$\Theta = \left\{ \lambda, \sigma, \delta, \alpha_w, \alpha_e, A_{e,0}, \nu_e^A, \psi(\cdot, 1995), \zeta_{w,0}, \zeta_{e,0}, \nu_w^\zeta, \nu_e^\zeta \right\}$$

The set of model implied moments are denoted by $\mathcal{M}^{\text{model}}(\Theta)$ as a function of Θ conditional on externally calibrated parameters. The corresponding set of moments which can be measured from the data is denoted by $\mathcal{M}^{\text{data}}$. Then, internally calibrated parameters, $\Theta^\star$ can be defined as the minimizer of the following objective function

$$\Theta^\star = \arg\min_{\Theta} \sum_{k=1}^{K} \omega_k \frac{\left| \mathcal{M}_k^{\text{model}}(\Theta) - \mathcal{M}_k^{\text{data}} \right|}{0.5 \left(\left| \mathcal{M}_k^{\text{model}}(\Theta) \right| + \left| \mathcal{M}_k^{\text{data}} \right| \right)}$$

where subscript k denotes a particular moment, and ω_k corresponds the weight assigned for that moment such that $\sum_k \omega_k = 1$. In the calibration, equal weights are used.

Results. Calibrated parameter values are depicted in Table 2.2 along with targeted moments, and their data and model counterparts. Overall fit of the model to the empirical targets is quite reasonable. One important result that stems from the calibration exercise is that R&D cost parameter for East is around 3.6 times that of the West. This is a huge gap and an important determinant of the long run discrepancy between East and West regions. Another result of the calibration is that East is estimated to have an aggregate productivity level of 28.5% of the West. This gap is estimated to close by half in approximately 14 years. The half-lifes of R&D subsidy parameters are 10 years for West, and 16 years for East, which gives confidence to the choice of 10 years for the parameter ν_e^T.

Innovation step size is estimated to be around 1.15, which is close to other findings in the literature. Estimation reveals that quality spillovers, which is measured by δ parameter, are quite strong. It is found that the probability of getting a quality spillover from the market leader is 16.3%. Akcigit and Ates [2023] finds an annual probability of 8.4% for the US

Panel A: Parameter estimates			Panel B: Moments		
Parameter	Value	Description	Moment	Model	Data
λ	1.148	Innovation step size	East relative GDP per worker in 1995	0.632	0.643
σ	105.7	Elasticity of substitution	East relative GDP per worker in 2015	0.790	0.784
δ	0.163	Rate of quality diffusion	East relative wage in 1995	0.750	0.737
α_w	1.586	West R&D cost scale	East relative wage in 2015	0.806	0.812
α_e	5.652	East R&D cost scale	East average log price gap in 2015	-0.213	-0.213
$A_{e,0}$	0.285	East efficiency in 1995	West average R&D between 1995-2015 (%)	1.638	1.638
$log(2)/\nu_e^A$	14.156 years	Half-life of efficiency	East average R&D between 1995-2015 (%)	0.085	0.085
$\zeta_{w,0}$	0.100	West R&D subsidy rate	West R&D subsidy in 1995 (%)	0.135	0.135
$log(2)/\nu_w^\zeta$	9.927 years	Half-life of West R&D subsidy rate	West R&D subsidy in 2015 (%)	0.044	0.044
$\zeta_{e,0}$	0.249	East R&D subsidy rate	East R&D subsidy in 1995 (%)	0.038	0.038
$log(2)/\nu_e^\zeta$	16.317 years	Half-life of East R&D subsidy rate	East R&D subsidy in 2015 (%)	0.005	0.005

Table 2.2: Internally calibrated parameters and model fit

economy for the period before 1980, where it declines to 3% until 2000s. Compared to these findings, German economy can be regarded as more dynamic than the US economy in terms of between-firm technology spillovers.

The elasticity of substitution is estimated to be quite high, approximately 105. It should be noted that, at face value, this finding suggests that West and East goods are strong substitutes of each other. Given that the price gap data consists of product categories at the 9-digit level, this result seems reasonable. Another important point on this finding is that demand elasticity is actually not constant across firms, it depends on relative sizes and market power of the firms. $\sigma = 105$ can be thought as an upper limit, while the elasticity that firms face lies between one and this value, which is actually equal to a weighted average of these two limits, weights being the market share of firms.

Estimated initial quality gap in 1995 suggests considerable quality differences between West and East products. Fraction of sectors with a Western quality lead is estimated to be 0.73 in 1995, whereas only one fifth of sectors are led by East firms. The remaining 7% of sectors are in neck-and-neck competition. When we analyze the changes of these fractions from 1995 to 2015, we realize the fact that East firms' performance on the quality margin was quite poor. In 2015, 62% of sectors have a Western quality leader in the calibrated model, while only 6% of sectors are led by East firms. Because of the high speed of quality

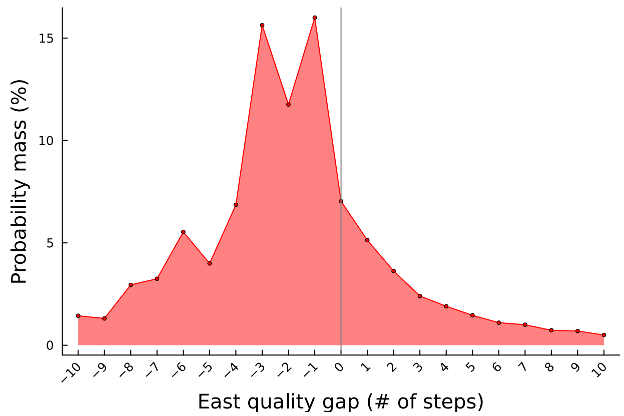
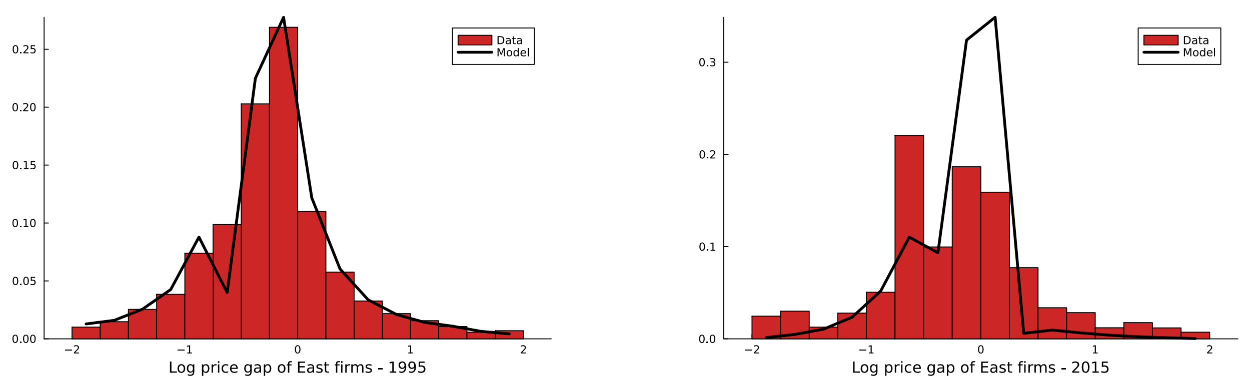


Figure 2.9: East quality gap distribution, 1995

diffusion, in 2015, 32% of sectors are in neck-and-neck competition. The full quality gap distribution between the steps -10 and 10 can be seen in Figure 2.9

Implied price gap distributions and their comparison with the data is depicted in Figure 2.10. Calibration mimics the distribution in 1995 quite closely. For 2015, the model implied distribution is accumulated more around zero although the average gap -0.213 is matched exactly.



(a) 1995

(b) 2015

Figure 2.10: Log price gap distributions: Model vs Data

Relative East efficiency is 28.5% of the West in 1995. Calibrated model implies that this level increased to 73.2% in 2015. Estimated path of relative efficiency of East labor force can be seen in Figure 2.11.

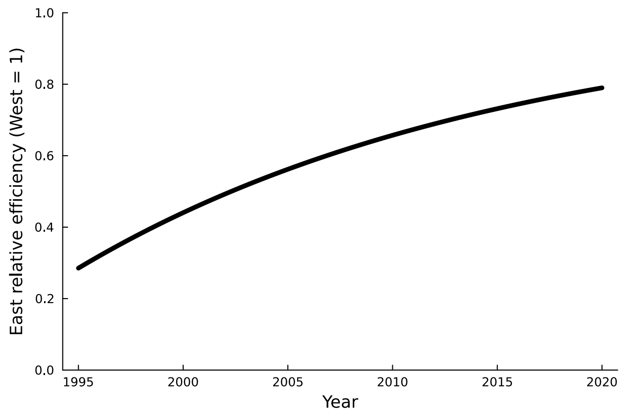


Figure 2.11: Relative production efficiency of East labor force, $A_e(t)$

2.5.2 Equilibrium properties

Mechanisms at play. In the calibrated economy, two main drivers determine the income convergence between the regions. The first one, the quality channel, captures the average level of technology (equivalently, product qualities), which evolve endogenously based on firms' forward-looking R&D decisions and knowledge spillovers. The second one, the efficiency channel, is an exogenous process by which East labor force learns from the frontier production techniques in the West and is disciplined empirically by the path of per capita income convergence between the two regions. Improvements in each of these margins help the average Eastern incomes approach Western levels. Next, we explore the dynamics of these margins in the calibrated model in order to assess their contributions to the evolution of relative incomes.

Figure 2.12 shows the changes in relative average quality of East firms, $\tilde{Q}_e(t) \equiv \frac{Q_e^{NT}(t)}{Q_w^{NT}(t)}$, and relative East production efficiency, $A_e(t)$. In the figure, 1995 values of both series are normalized to one for comparability of the two series with each other. Average quality of tradable good in a region is equal to $Q_i^{NT}(t)$, which also denotes the aggregate quality of nontradable goods in the region. The reason is the specific assumption given by (2.2).

As can be seen from Figure 2.12, the increase in East relative aggregate productivity is much larger than the increase in relative product qualities of East firms. In particular,

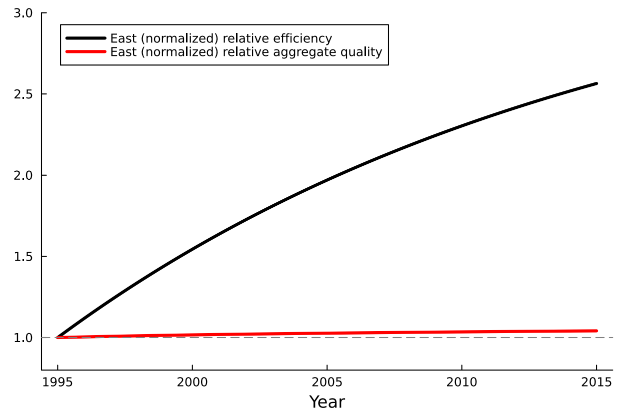


Figure 2.12: Comparison of channels, efficiency and quality

Note: The black line represents $A_e(t)$ where $A_e(1995)$ is normalized to one. Therefore, the y-axis gives the value of $A_e(t)$ relative to its initial value $A_e(1995)$. Similarly, the red line represents $\tilde{Q}_e(t) \equiv \frac{Q_e^{NT}(t)}{Q_w^{NT}(t)}$ where its value in 1995 is normalized to one.

relative East production efficiency rises more than 2.5 times its initial value in 2015. However, the increase in East product qualities relative to West qualities is muted, and its value in 2015 is slightly above its initial value 20 years ago. This finding suggests that East firms were not able to perform well in the product market competition with Western firms.

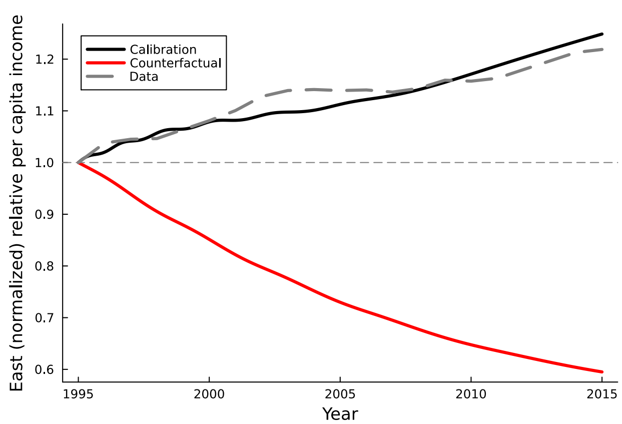


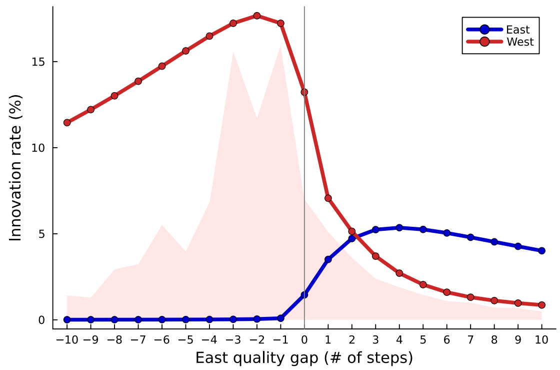
Figure 2.13: Counterfactual equilibrium without productivity spillovers

Note: Calibrated evolution of relative East per capita income is represented by black line. The red line which is labeled as counterfactual is the equilibrium evolution of the same series under the assumption that productivity spillovers are shut down. In particular, it is taken $A_e(t) = A_e(1995)$ for all t.

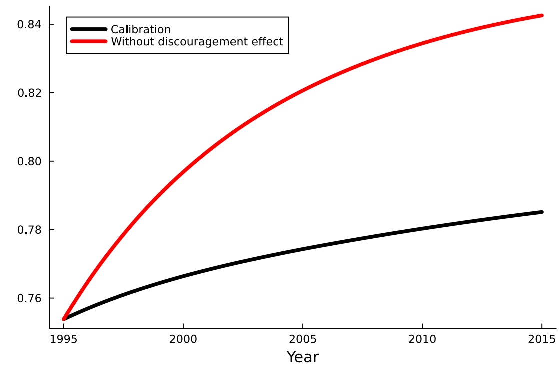
In order to evaluate the importance of productivity spillovers from West to East, we solve a counterfactual equilibrium in which the production efficiency level of East is kept its calibrated initial value, i.e. $A_e(t) = A_e(1995) \approx 0.285$. Figure 2.13 suggests the stark difference of our convergence measure, per capita income of East relative to West, between the counterfactual model and calibrated model. Without productivity spillovers, East households would have been even worse off over time in terms of purchasing power relative to the West.

Discouragement effect. As suggested by Figure 2.12, East firms have not improved their qualities fast enough relative to West. Further inspection suggests that the reason why Eastern firms do not improve their product quality faster is the lack of sufficient competitive forces that incentivize them to do so. At the time of unification, most Eastern firms are technologically behind their Western rivals and many by some large distance, as shown in Figure 2.9. As a result, their hope to upgrade their technology sufficiently enough to catch up and surpass their rivals is slim discouraging them from investing in R&D. To show this, we plot average innovation rates chosen by firms in calibration in Figure 2.14a. The shaded area shows the calibrated initial quality gap distribution. East firms reduce their innovation efforts radically as they begin falling behind their competitors. Competition forces that incentivize East firms to exert more innovation starts kicking in as East firms have positive quality gaps. However, the mass of such East firms is calibrated to be too low—most of East firms are behind Western firms, and they invest in R&D very little.

This result is a reflection of the key role strategic forces play in driving firms' endogenous forward-looking decisions in step-by-step innovation models. Firms intensify innovation effort when in close competition with their rivals for market leadership, but they reduce it when they face less competition, which occurs at wider technology gaps between rivals. In fact, comparing the dynamics in the baseline economy to a counterfactual one in which Eastern firms are assumed to exert the maximum effort at every point in time can give an idea about the magnitude of the discouragement effect. Figure 2.14b demonstrates that the impact of this discouragement effect is substantial. In the counterfactual economy, the Eastern firms indeed innovate more intensively—enough to ensure that the technological gap with the West closes faster over time. Consequently, the average quality level in the East relative to that



(a) Average innovation rates



(b) Relative aggregate East quality

Figure 2.14: Discouragement effect

Note: Shaded area in panel (a) shows the calibrated quality gap distribution in 1995. In panel (b), the red line shows the counterfactual evolution of relative aggregate East quality, $Q_e^{NT}(t)$, had all of East firms performed the highest innovation rate at the time in their region. Without GE effects.

in the West (red line) rises 6 percentage points in 2015, in contrast to the muted increase observed in the calibrated baseline economy (black line).

2.6 Policy analysis

In this section, we evaluate several alternative policy settings in terms of their implications regarding the technological catch-up between the two German regions and consumer welfare. We first explore whether delaying the unification process could have potentially dampened the discouragement effect that hurt Eastern firms' incentives. Then, we analyze alternative subsidies—R&D subsidies, production subsidies, and outright consumption support—that were considered or part of the policy kit back at the time. Finally, we entertain a policy that would encourage technology flows from the frontier Western firms to the East via licensing. To be sure, our goal here is not to evaluate actual policies adopted in Germany back at the time; rather, it is to comprehend the implications of various policy options in order to assess best practices that could have sped up technological and income convergence between the East and West.

Welfare. The measure of welfare is in consumption equivalent terms. Let $\mathcal{U}_i(T)$ denote the total discounted utility of the representative household of region i over a period of length T

$$\mathcal{U}_i(T) \equiv \int_{T_0}^{T} e^{-\rho(t-T_0)} \log \{Q_i(t)C_i(t)\}\, dt$$

Let $\mathcal{U}_i^{\text{cal}}(T)$ denote this value under the benchmark calibration. If the consumption equivalent welfare change with a time horizon of T associated with a certain policy is denoted by $\gamma_i(T)$, then $\gamma_i(T)$ satisfies the equation below

$$\int_{T_0}^{T} e^{-\rho(t-T_0)} \log \left\{ Q_i^{\text{policy}}(t) C_i^{\text{policy}}(t) \left[1-\gamma_i(T)\right] \right\} dt = \mathcal{U}_i^{\text{cal}}(T)$$

where $Q_i^{\text{policy}}(t)$ and $C_i^{\text{policy}}(t)$ refer to the aggregate quality of consumed goods and their quantity under the new policy regime, respectively. A positive $\gamma_i(T)$ means that the consumers in region i are better off under the new policy with a $(100 \cdot \gamma_i(T))\,\%$ consumption equivalent welfare increase over a horizon of T. Conversely, a negative $\gamma_i(T)$ implies $|(100 \cdot \gamma_i(T))|\,\%$ decline in welfare.

2.6.1 Delayed unification

Given the discouragement effect on Eastern firms exerted by Western ones in the post-unification economy—as examined in Section 2.5.2—a sensible follow-on question is if delaying unification could have helped Eastern firms and incomes. Indeed, some policy makers at the time pushed for a delayed or phased integration, a perspective that is still debated today. We now assess this idea in a counterfactual economy, in which excessively high tariff rates imposed by the central German government on East imports from the West, which effectively prohibits the access of Western goods into the East market in early years while protecting Eastern firms before declining gradually.

Figure 2.15 shows the path of regional average qualities and changes in consumption-

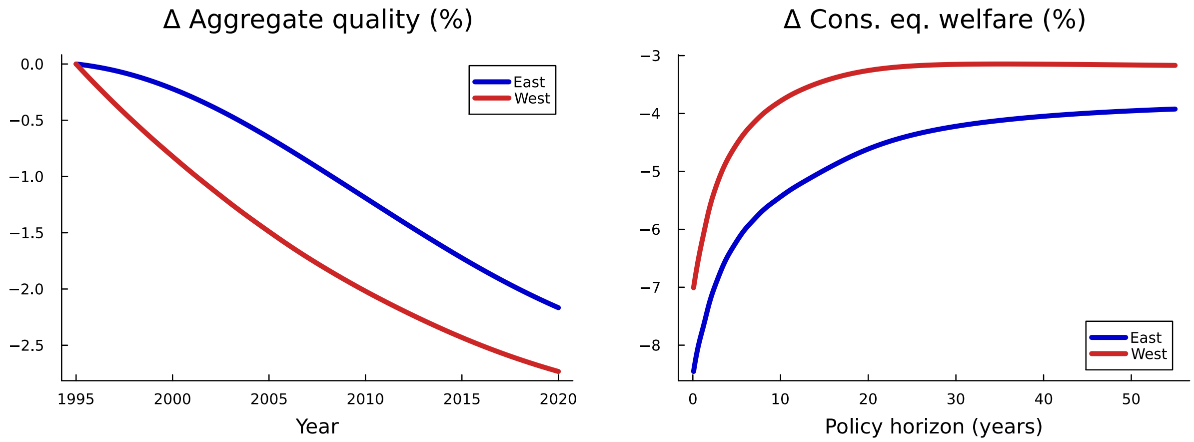


Figure 2.15: Effects of delayed unification

Notes: The left panel shows the change in aggregate quality of a region's products over time relative to the benchmark calibration. The right panel shows consumption equivalent welfare change relative to calibration over different time horizons.

equivalent welfare over different horizons. The left panel of the figure indicates that regional average qualities are lower than their baseline paths, and the losses grow over time. While the dynamic loss is larger in the West in proportional terms; in relative terms, the Eastern average technology is closer to the West in the delayed unification scenario when compared with the baseline. This seemingly improved Eastern technological position relative to the West clearly occurs at the expense of overall technological deterioration. As shown in the right panel of Figure 2.15, these dynamics lead to substantial welfare losses in both regions.

The main reason for dynamic losses in average quality is the reduced innovation incentives of firms caused by higher trade barriers. While higher trade costs on Western imports shrink the market size from sales abroad for Western firms, they also reduce the competitive pressure on many technologically laggard Eastern firms, with their domestic market being protected by tariffs. Moreover, while tariffs are raised unilaterally only on the Western firms, the demand from the West for Eastern products also diminishes reciprocally—recall that trade is balanced—hurting Eastern exporters' profits and their innovation drive, as an indirect effect. Delaying unification, and hurting innovation incentives of the technologically advanced Western firms is a bad idea also for another reason. Recall that benchmark calibration suggests sizable technology spillovers from market leaders to followers. Because, many

Eastern firms are laggard, they are actually main beneficiary of these spillovers. Therefore, as the frontier technology, which is represented by Western firms, begins investing less in R&D due to reduced market size effect, the laggard region also loses in the long run because of less effective spillovers in the future. In sum, delayed unification depresses overall technological development, weighing on consumer welfare.[17]

In light of these findings, an alternative policy design could be providing support to R&D in the East in conjunction with delaying the unification to withstand its negative effects on innovation incentives. To entertain this alternative, we consider the same path of increased import tariffs as above, but with the tariff revenue collected in the East being used to subsidize R&D activity of Eastern firms. Figure 2.16 compares the implications for the average quality and welfare in the East with the previous case—delayed unification without R&D subsidy. As can be expected, the average quality path is higher with the help of additional support to R&D, even to the extent that the path is better than in the calibrated baseline economy (note that the change is relative to the baseline, and the dashed line is in the positive territory). However, this policy still implies lower welfare in the East relative to the baseline for about half a century. While the R&D support helps correcting for the negative effects on innovation of trade barriers and make Eastern consumers better off in the longer run, they are indeed worse off in the short run compared with the delayed unification scenario. The reason is that R&D subsidy to East causes labor reallocation from production sector to research sector reducing output in the East, whereas the benefits of the subsidy on product qualities materialize gradually over time.

17. The reason why the magnitude of the welfare loss in the West is smaller than in the East despite the larger dynamic losses in average quality reflects the fact that the size of the regional Western market is relatively much larger than the Eastern market for exports. As such, the income loss from reduced exports is relatively less detrimental in welfare terms.

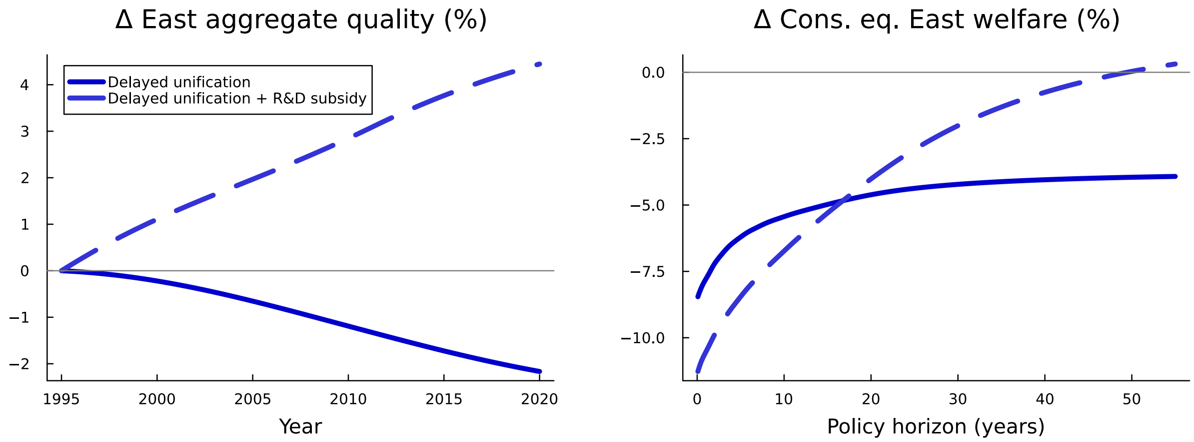


Figure 2.16: Effects of delayed unification with R&D subsidy to East

Notes: Figures compare effects of delayed unification policies with and without R&D subsidies to East. This subsidy is financed by the tariff revenue collected. The left panel shows the change in aggregate quality of East products over time relative to the benchmark calibration. The right panel shows consumption equivalent welfare change relative to calibration over different time horizons.

2.6.2 Subsidy policies

Subsidies to the East. Following the reunification, the German government established several schemes to lift up the Eastern economy, including direct transfers to households and producers in the East funded by taxpayers living in the more affluent West German states. In this part, we evaluate how various transfer policies could be used to support Eastern economy in the aftermath of the unification. We start by comparing Western support through (i) R&D subsidies to Eastern firms, (ii) production subsidy to Eastern firms, and (iii) lump-sum transfers to Eastern consumers. In each case, the change in transfers last forever and amount to 0.3 percent of the German GDP every year.

As shown in Figure 2.17, consumption and production subsidies lead to a slight decline in the average quality of Eastern products relative to the baseline. That said, both policies are effective in boosting the welfare of Eastern households, as shown in Figure 2.18. Not surprisingly, consumption subsidies provide windfall gains to the consumers in the East, while production subsidies lower the effective cost of production in the East, supporting production, income and consumption in that region.[18]

18. The increased demand for production workers and higher wages also raise the cost of doing R&D in

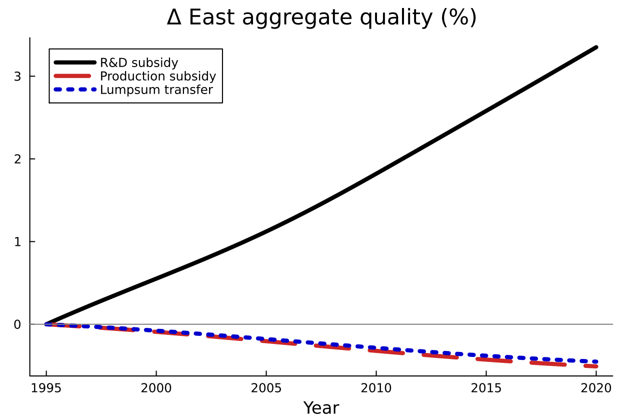


Figure 2.17: Effects of subsidies to East on East technology

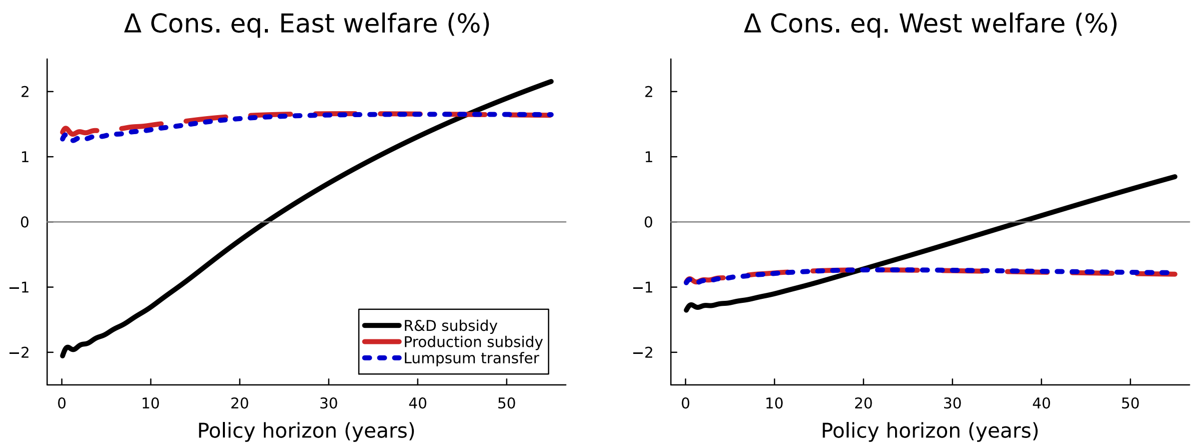


Figure 2.18: Effects of subsidies to East on welfare

Notes: Left panel and right panel show the welfare effect of subsidies on East and West consumers, respectively, over different time horizons.

A plausible question is then whether the same amount of resources used in the previous subsidy counterfactuals could be spared for R&D subsidies, potentially boosting technological convergence and raising welfare. Figures 2.17 and 2.18 confirm this conjecture, though with some important nuances. To start, R&D subsidies prove very effective in supporting technological convergence, improving the average quality path notably relative to the baseline (the solid black line in Figure 2.17). However, this acceleration in product qualities does not immediately translate into welfare gains. Indeed, R&D subsidies create welfare losses

the production-subsidy counterfactual. This is the reason why the relative quality path in this experiment (the red dashed line in Figure 2.17) is slightly worse than in the baseline.

in the East for horizons up to two decades, and it takes about four decades for the Eastern consumers to enjoy similar gains as in the case of consumption and production subsidies. The reason is that substantial R&D support raise labor demand for this activity, putting pressure on wages and production workforce, eroding Eastern firms competitiveness vis-à-vis their Western counterparts. Therefore, these gains materialize with considerable lags, with production or consumption subsidies turning out to be more beneficial options when the policy maker's horizon is shorter than four decades.

Turning to the implications of these alternative policies for Western consumers, right panel of Figure 2.18 shows that all these schemes to provide additional support to the East reduce welfare in the West, except in the long run in the case of R&D subsidies. The benefits from improved product quality eventually dominate the welfare changes for the Western consumer, making them better off only when a horizon of at least four decades is considered.

R&D subsidy in the West. This investigation reveals that among the analyzed subsidy schemes, only R&D subsidies bolster technological advancement in the East, though at the expense of consumers' welfare in both regions. An alternative scheme could be using the same amount of resources to support R&D in the West, given that Western firms are more efficient in doing R&D—the cost of obtaining the same rate of innovation is lower in the West—and the East could still benefit from faster technological improvements through knowledge spillovers. Indeed, as shown in Figure 2.19, Eastern average quality quickly benefits from this policy via spillovers, even surpassing the path that emerges in the case of increased R&D transfer to the East in about two decades (and without much loss prior to that). In addition, welfare in the both regions improves (except for the immediate horizons in the West) owing to improved quality of products sourced from both regions. Welfare effects of this policy is demonstrated in Figure 2.20.

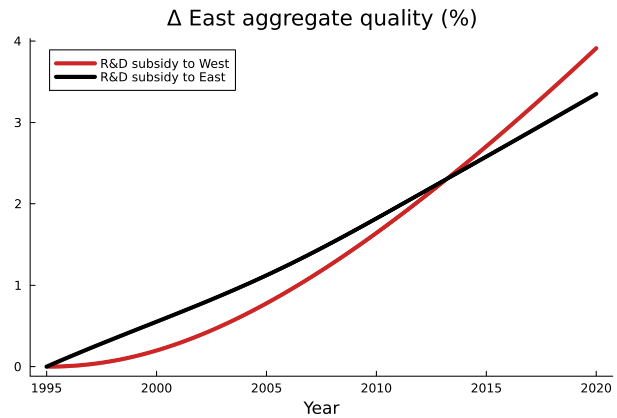


Figure 2.19: Effects of R&D subsidy to West on East technology

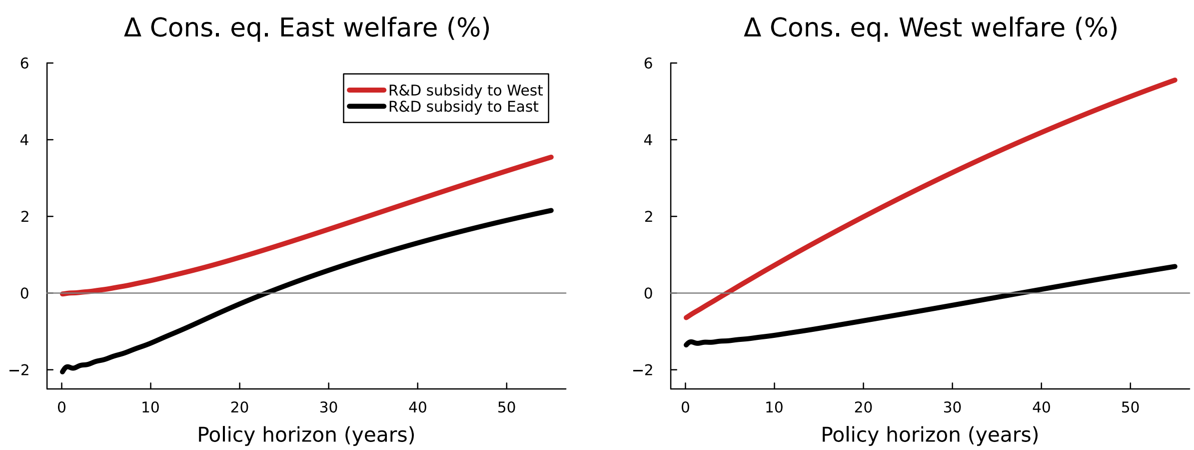


Figure 2.20: Effects of R&D subsidy to West on welfare

Notes: Left panel and right panel show the welfare effect of subsidies on East and West consumers, respectively, over different time horizons.

2.6.3 Technology transfer from West

Next, we consider an alternative policy to directly alter the dynamics of technological convergence; that is, having West German firms to share their frontier technologies with laggard Eastern firms in their sectors. The thought experiment is that at the time of integration, Western frontier firms transfer their technology to Eastern competitors via licensing, though it is unlikely that it can fully be absorbed by Eastern producers. As a result of this policy, Eastern firms enter the integrated economy at a better position than they would in the calibrated economy, such that, on average, their initial position is closer to the Western firms by a certain number of technology steps.

To be able to draw comparisons to the previous exercises, we determine the magnitude of technology transfer in a way that the cost of this policy on the Western side is exactly equal to the previous cases. The cost of this policy on Western firms is the loss in future profit stream as a result of closer competition with East, translating into lower firm valuations in 1995 in net present discounted terms. We then solve for the number of technology steps to be transferred to East competitors in order to match the loss in Western firm valuations to the discounted present value of West household taxation of 0.3% of German GDP every year, which is the life-time cost of previous subsidy schemes on West households. The idea is that instead of West households pay for the cost of subsidies of previous section in every year, in technology transfer policy, they pay it once only in 1995 via a decline in their financial portfolio. [19]

Technology transfer has an immediate large impact on the relative technology level of Eastern firms, as shown in Figure 2.21. Given that the support is assumed to be one time at the onset of integration, the relative technology level of Eastern firms starts to regress after

19. We do not change a East firm's position if it already had a positive gap to its competitor in 1995. Moreover, we do not allow follower East firms to leapfrog West firms after the technology transfer. At most, they can become neck-and-neck with their competitors. This situation arises if the initial gap between East and West firms in a sector is less than the number of steps of the technology transfer.

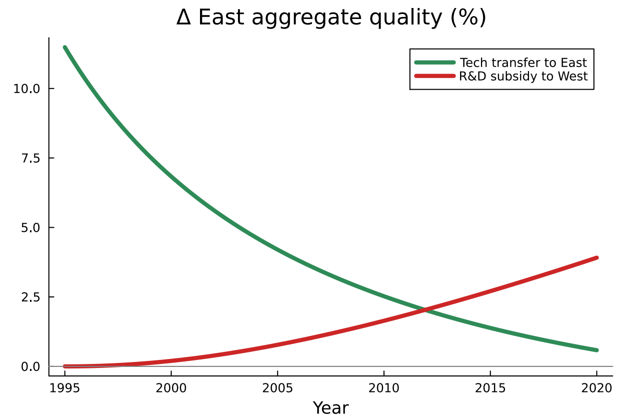


Figure 2.21: Effects of technology transfer to East on East technology

the initial jump, though always remaining at levels materially higher than the path arising in the calibrated economy. Secondly, the technologically improved position of Eastern firms translate into higher market shares, incomes, and considerable permanent welfare gains, especially over shorter horizons. In addition, welfare in the West also improves in the short term owing to the expanded consumption of better-quality goods with Eastern products that benefit from technology transfer, a distinguishing implication of technology transfer compared to previously analyzed subsidies (Figure 2.22). The loss of income to Eastern firms becomes the dominant force when longer horizons are considered, leading to small declines in the welfare of Western consumers. Hence, when compared with the scenario of increased R&D subsidies to the Western firms, this policy appears to shuffle welfare gains across time and space: toward earlier periods and Eastern consumers.

2.6.4 Policy mix

Is there a way to achieve a more balanced welfare improvement across time and space mixing the most effective policies—subsidizing R&D of Western firms and initial transfer of frontier Western technology to laggard East? In this section, we examine this policy mix, with the total resources allocated remaining the same as in previous experiments and half of it being spent on each policy. Figure 2.23 demonstrates the Eastern average quality under this

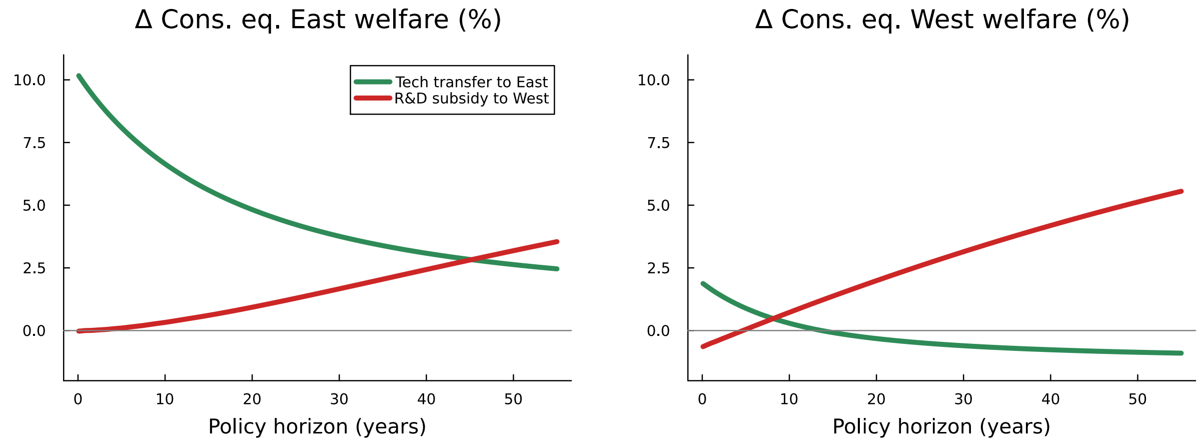


Figure 2.22: Effects of technology transfer to East on welfare

Notes: Left panel and right panel show the welfare effect of subsidies on East and West consumers, respectively, over different time horizons.

scenario. As could be expected, the path lies in between the ones obtained separately in each scenario, implying still a notable improvement relative to the baseline. The welfare gains also follow similar patterns, with Western consumers now observing an improvement relative to the baseline over any horizon, though at a lower rate than that in the case of subsidizing Western R&D alone. (Figure 2.24)

To sum, our results indicate that it would be possible to accelerate technological advancement in the East in a way that is beneficial to all consumers over any policy horizon by having frontier Western technologies to be shared with laggard Eastern firms initially, and making up for potential losses in the West via subsidies to Western R&D, as West is the most productive region in Germany in terms of research and innovation.

2.7 Conclusion

In this study, we attempt to contribute to the perennial debate on the persistent income and technology gaps between the two regions of Germany—the West and the East. Motivated by similar observations at granular product-level data, we propose a theory that links firms' endogenous innovation incentives, which drive aggregate dynamics of technological develop-

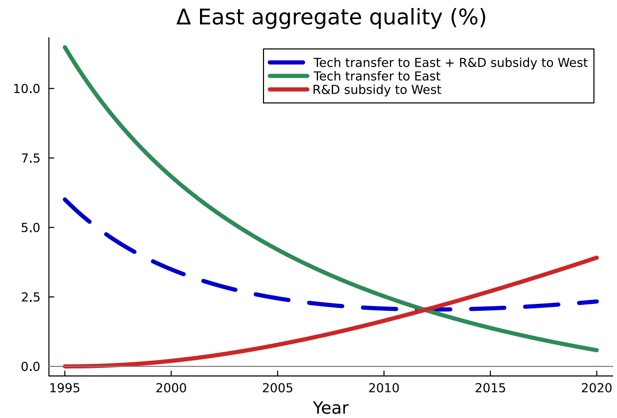


Figure 2.23: Effects of the policy mix on East technology

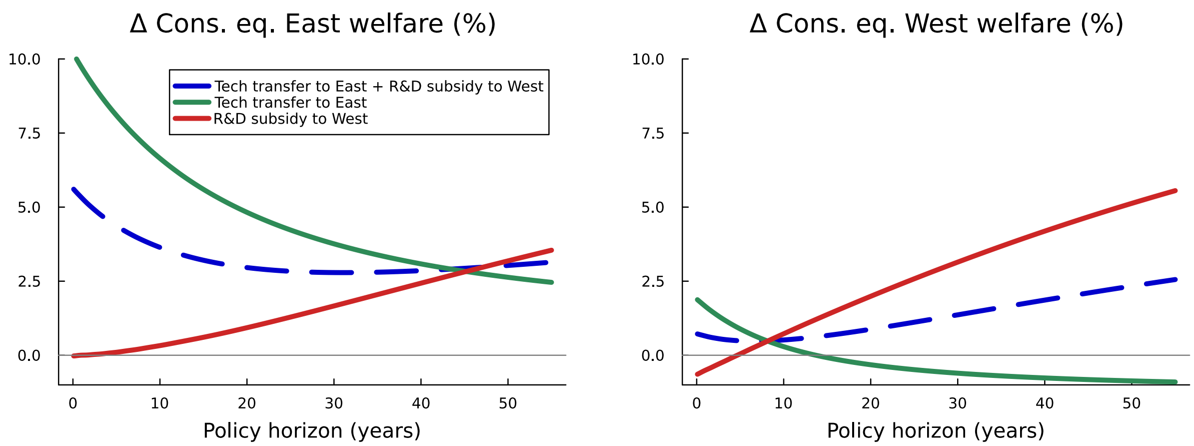


Figure 2.24: Effects of the policy mix on welfare

Notes: Left panel and right panel show the welfare effect of subsidies on East and West consumers, respectively, over different time horizons.

ment and income, to their positions in market competition, which, in turn, reflects the level of technology they command relative to their rivals. We quantify the key features of this model—notably, the dynamics of relative product qualities—relying on a novel dataset of firm-level prices over narrowly defined products.

The quantitative analysis provides various insights regarding technological convergence between the two regions and policy implications. To start, the results point to a novel technological factor coupled with the dynamics of market competition as the root cause of persistent product quality disparities between the East and the West and the slow pace of economic convergence. Policies that aim at boosting technology upgrading of Eastern firms could alleviate this problem. That said, results emphasize that standard support schemes such as R&D subsidies to Eastern firms would not be welfare-improving over the relevant policy horizons in light of the estimated high cost of doing R&D in the East. More effective use of resources would need to rely on technology transfers from Western firms via licensing while supporting their innovative activity, respecting their comparative advantage and ensuring the flow of knowledge spillovers.

Certainly, the analysis bears broader relevance to topics concerning economic development. A central theme in macroeconomics is the design of appropriate policies to support the convergence of developing economies to high-income levels. The fact that economic convergence could drag for decades even among two regions that share a common language, common cultural backgrounds, and common borders only adds to the intricacies of economic development and optimal design of industrial policies to support it. This study could potentially shed light on underexplored technological factors that could perpetuate and impinge upon economic convergence and the most effective policies to overcome these challenges.

REFERENCES

Daron Acemoglu and Ufuk Akcigit. Intellectual property rights policy, competition and innovation. *Journal of the European Economic Association*, 10(1):1–42, 2012.

Daron Acemoglu, Philippe Aghion, and Fabrizio Zilibotti. Distance to Frontier, Selection, and Economic Growth. *Journal of the European Economic Association*, 4(1):37–74, 2006.

Daron Acemoglu, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr. Innovation, Reallocation, and Growth. *American Economic Review*, 108(11):3450–3491, 2018.

Philippe Aghion and Peter Howitt. A model of growth through creative destruction. *Econometrica: Journal of the Econometric Society*, pages 323–351, 1992.

Philippe Aghion, Christopher Harris, Peter Howitt, and John Vickers. Competition, imitation and growth with step-by-step innovation. *Review of Economic Studies*, 68(3):467–492, 2001.

Philippe Aghion, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. Competition and Innovation: An Inverted-U Relationship. *Quarterly Journal of Economics*, 120 (2):701–728, 2005.

Philippe Aghion, Jing Cai, Mathias Dewatripont, Luosha Du, Ann Harrison, and Patrick Legros. Industrial policy and competition. *American Economic Journal: Macroeconomics*, 7(4):1–32, 2015.

Ufuk Akcigit and Sina T. Ates. Ten facts on declining business dynamism and lessons from endogenous growth theory. *American Economic Journal: Macroeconomics*, 13(1):257–98, 2021.

Ufuk Akcigit and Sina T Ates. What happened to us business dynamism? *Journal of Political Economy*, 131(8):2059–2124, 2023.

Ufuk Akcigit and William R Kerr. Growth through heterogeneous innovations. *Journal of Political Economy*, 126(4):1374–1443, 2018.

Ufuk Akcigit, Sina T. Ates, and Giammario Impullitti. Innovation and trade policy in a globalized world. 2018. National Bureau of Economic Research Working Paper 24543.

Ufuk Akcigit, Douglas Hanley, and Stefanie Stantcheva. Optimal Taxation and R&D Policies. *Econometrica*, 90(2):645–684, 2022.

Ufuk Akcigit, Harun Alp, André Diegmann, and Nicolas Serrano-Velarde. Committing to grow: Size-dependent regulations and firm dynamics in east germany. Technical report, 2023.

Ufuk Akcigit, Sina T Ates, Josh Lerner, Richard R Townsend, and Yulia Zhestkova. Fencing off silicon valley: Cross-border venture capital and technology spillovers. *Journal of Monetary Economics*, 141:14–39, 2024.

George A Akerlof, Andrew K Rose, Janet L Yellen, Helga Hessenius, Rudiger Dornbusch, and Manuel Guitian. East germany in from the cold: the economic aftermath of currency union. *Brookings Papers on Economic Activity*, 1991(1):1–105, 1991.

Andrew Atkeson and Ariel Burstein. Aggregate implications of innovation policy. *Journal of Political Economy*, 127(6):2625–2683, 2019.

Robert J. Barro. Convergence and Modernization Revisited. 2012. National Bureau of Economic Research Working Paper 18295.

Robert J. Barro and Xavier Sala-i Martin. Convergence. *Journal of Political Economy*, 100 (2):223–251, 1992.

Sascha O. Becker, Lukas Mergele, and Ludger Woessmann. The separation and reunification of germany: Rethinking a natural experiment interpretation of the enduring effects of communism. *Journal of Economic Perspectives*, 34(2):143–171, 2020.

Olivier Jean Blanchard and Lawrence F. Katz. Regional Evolutions. *Brookings Papers on Economic Activity*, 23(1):1–76, 1992.

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

Michael C. Burda. Factor reallocation in eastern germany after reunification. *American Economic Review*, 96(2):368–374, 2006.

Michael C. Burda. 30 jahre deutsche einheit: Wie steht es wirklich? *Wirtschaftsdienst*, 100 (6):390–391, 2020.

Michael C Burda and Jennifer Hunt. From reunification to economic integration: Productivity and the labor market in eastern germany. *Brookings papers on economic activity*, 2001(2):1–92, 2001.

Ricardo J Caballero and Adam B Jaffe. How high are the giants' shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. *NBER macroeconomics annual*, 8:15–74, 1993.

Jie Cai, Nan Li, and Ana Maria Santacreu. Knowledge diffusion, trade, and innovation across countries and sectors. *American Economic Journal: Macroeconomics*, 14(1):104–145, 2022.

Fabio Canova and Morten Ravn. The Macroeconomic Effects of German Unification: Real Adjustments and the Welfare State. *Review of Economic Dynamics*, 3(3):423–460, 2000.

Gerald Carlino and William R Kerr. Agglomeration and innovation. *Handbook of regional and urban economics*, 5:349–404, 2015.

Christophe Carrincazeaux, Yannick Lung, and Alain Rallet. Proximity and localisation of corporate r&d activities. *Research Policy*, 30(5):777–789, 2001.

Francesco Caselli and Wilbur John Coleman II. The U.S. Structural Transformation and Regional Convergence: A Reinterpretation. *Journal of Political Economy*, 109(3):584–616, 2001.

Laurent Cavenaile, Murat Alp Celik, and Xu Tian. Are markups too high? competition, strategic innovation, and industry dynamics. *mimeo*, 2023.

Craig Chikis, Jonathan E. Goldberg, and J. David López-Salido. Do Low Interest Rates Harm Innovation, Competition, and Productivity Growth? 2021. CEPR Discussion Papers 16184.

Jaedo Choi and Younghun Shim. From Adoption to Innovation: State-Dependent Technology Policy in Developing Countries. 2024. Structural Transformation and Economic Growth Working Paper 091.

Irwin L. Collier. On the first year of german monetary, economic and social union. *Journal of Economic Perspectives*, 5(4):179–186, December 1991. doi:10.1257/jep.5.4.179. URL `https://www.aeaweb.org/articles?id=10.1257/jep.5.4.179`.

Chiara Criscuolo, Ralf Martin, Henry G. Overman, and John Van Reenen. Some causal effects of an industrial policy. *American Economic Review*, 109(1):48–85, 2019.

Klaus Desmet, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg. The geography of development. *Journal of Political Economy*, 126(3):903–983, 2018.

Euractiv. Germany's reunification, 33 years on: East-west divide persists. September 29 2023.

Sebastian Findeisen, Sang Yoon Lee, Tommaso Porzio, and Wolfgang Dauth. Transforming Institutions: Labor Reallocation and Wage Growth in a Reunified Germany. 2021. mimeo.

Nicola Fuchs-Schündeln and Rima Izem. Explaining the low labor productivity in East Germany – A spatial analysis. *Journal of Comparative Economics*, 40(1):1–21, 2012.

Nicola Gennaioli, Rafael La Porta, Florencio Lopez De Silanes, and Andrei Shleifer. Growth in regions. *Journal of Economic Growth*, 19(3):259–309, 2014.

Bronwyn H Hall, Adam B Jaffe, and Manuel Trajtenberg. The nber patent citation data file: Lessons, insights and methodological tools, 2001.

Sebastian Heise and Tommaso Porzio. Transforming Institutions: Labor Reallocation and Wage Growth in a Reunified Germany. 2022. National Bureau of Economic Research Working Paper 30298.

Jennifer Hunt. Staunching Emigration from East Germany: Age and the Determinants of Migration. *Journal of the European Economic Association*, 4(5):1014–1037, 2006.

Adam B Jaffe, Manuel Trajtenberg, and Rebecca Henderson. Geographic localization of knowledge spillovers as evidenced by patent citations. *the Quarterly journal of Economics*, 108(3):577–598, 1993.

Adam B Jaffe, Manuel Trajtenberg, and Michael S Fogarty. Knowledge spillovers and patent citations: Evidence from a survey of inventors. *American Economic Review*, 90(2):215–218, 2000.

Charles I Jones. R & d-based models of economic growth. *Journal of political Economy*, 103 (4):759–784, 1995.

Réka Juhász, Nathan J. Lane, and Dani Rodrik. The new economics of industrial policy. 2023. National Bureau of Economic Research Working Paper 31538.

Tor Jakob Klette and Samuel Kortum. Innovating firms and aggregate innovation. *Journal of political economy*, 112(5):986–1018, 2004.

Nathan Lane. Manufacturing Revolutions: Industrial Policy and Industrialization in South Korea. 2022. University of Oxford, mimeo.

Ernest Liu. Industrial Policies in Production Networks. *The Quarterly Journal of Economics*, 134(4):1883–1948, 2019.

Ernest Liu, Atif Mian, and Amir Sufi. Low interest rates, market power, and productivity growth. *Econometrica*, 90(1):193–221, 2022.

Alfred Marshall. *Principles of economics: unabridged eighth edition*. Cosimo, Inc., 2009.

Matthias Mertens and Steffen Müller. The east-west german gap in revenue productivity: Just a tale of output prices? *Journal of Comparative Economics*, 50(3):815–831, 2022.

Enrico Moretti. The effect of high-tech clusters on the productivity of top inventors. *American Economic Review*, 111(10):3328–3375, 2021.

Karl-Heinz Paqué. *Die Bilanz. Eine wirtschaftliche Analyse der Deutschen Einheit*. Hanser, Munich, 2009.

Politico. German election exposes lingering east-west political divide. September 24 2021.

Xavier X. Sala-i-Martin. Regional cohesion: Evidence and theories of regional growth and convergence. *European Economic Review*, 40(6):1325–1352, 1996.

Hans-Werner Sinn. Germany's Economic Unification: An Assessment after Ten Years. *Review of International Economics*, 10(1):113–128, 2002.

Dennis J. Snower and Christian Merkl. The caring hand that cripples: The east german labor market after reunification. *American Economic Review*, 96(2):375–382, 2006.

The Economist. Special report: A survey of germany. February 1999.

The Economist. Still not over the wall. November 2014a.

The Economist. Twenty-five years on. November 2014b.

The Economist. A generation after germany reunited, deep divisions remain. May 2023.

Harald Uhlig. Regional labor markets, network externalities and migration: The case of german reunification. *American Economic Review*, 96(2):383–387, 2006.

Harald Uhlig. The slow decline of east germany. *Journal of Comparative Economics*, 36(4): 517–541, 2008.

Manès Weisskircher. The strength of far-right afd in eastern germany: The east-west divide and the multiple causes behind 'populism'. *The Political Quarterly*, 91(3):614–622, 2020.

APPENDICES

A Derivation of agent HJB equation

Let $\mathcal{U}_i^T(\varepsilon, t)$ denote the life-time utility of type-T agents living in location i, conditional on taste ε. For an infinitesimal time interval of $dt > 0$, the HJB equation can be written in discrete time as follows

$$\begin{aligned}\mathcal{U}_i^T(\varepsilon, t) &= A_i \varepsilon C_i^T(t) dt \\ &+ \frac{1}{1+\rho dt}\left[\zeta dt \cdot \int \left(\max_j \mathcal{U}_j^T(e_j, t+dt)\right) f_\epsilon(\mathbf{e}) d\mathbf{e} + (1-\zeta dt)\cdot \mathcal{U}_i^T(\varepsilon, t+dt)\right] \quad (26)\end{aligned}$$

The first term in the right hand side of equation (26) represents the utility flow due to consumption, amenities and location taste. The second term represents continuation value discounted by time preference parameter ρ. This continuation value is the expected value of drawing a migration shock whose rate is ζ. If agent updates location preferences, then she migrates to the best location for herself in terms of the discounted sum of utility. Otherwise, she stays in the same location.

Using notation defined by equation (1.5), we can organize the terms in (26) and take limit $dt \to 0$ to derive the HJB equation in continuous time given by (1.6) in the main text.

B Proof of Theorem 1.3.1

We start by taking first order condition of the maximization problem inside (1.9). It reads as

$$W_i^R(t)\left(1-s_i\right) = \frac{1}{\theta}\alpha_i(t)^{\frac{1}{\theta}} R_i(n,t)^{\frac{1}{\theta}-1} n^{1-\frac{1}{\theta}}\left[\mathcal{V}_i(n+1,t) - \mathcal{V}_i(n,t)\right]$$

Using notation of inventor employment per product line, $r_i(n,t)$, first order condition can be rewritten as

$$W_i^R(t)\,(1-s_i) = \alpha_i(t)^{\frac{1}{\theta}}\frac{1}{\theta}r_i(n,t)^{\frac{1}{\theta}-1}\left[\mathcal{V}_i(n+1,t) - \mathcal{V}_i(n,t)\right] \tag{27}$$

I conjecture that the solution to HJB equation (1.9) is $\mathcal{V}_i(n,t) = nv_i(t)Y(t)$ for a function $v_i(t)$. Replacing this conjecture into (27) implies that per product line inventor employment is independent of the number of product lines the firm owns, n, such that $r_i(n,t) = r_i(t)$, and equation (1.11) can be derived. Per product line innovation rate can be derived from R&D production function (1.8) substituting $R_i(n,t) = r_i(t)n$. This gives us equation (1.10) in Theorem 1.3.1 which states that per product line innovation rate is independent of the firm size proxied by n.

Finally replacing conjecture and first order condition (27) into the HJB equation (1.9), and denoting the growth rate of aggregate output as $g(t) \equiv \frac{\dot{Y}(t)}{Y(t)}$, equation (1.12) is derived. This equation governs the time evolution of $v_i(t)$ and states that it is independent of n, as conjectured.

C Proof of Theorem 1.3.2

First order condition to the maximization problem (1.14), after substituting $\mathcal{V}_i(1,t) = v_i(t)Y(t)$, can be found as below

$$W_i^R(t)(1-s_i) = \frac{1}{f}\frac{1}{\theta}\alpha_i(t)^{\frac{1}{\theta}}\tilde{r}_i(t)^{\frac{1}{\theta}-1}v_i(t)Y(t) \tag{28}$$

Substituting (28) into equation (1.14) results in (1.15) as stated by the Theorem. Finally we combine incumbent and entrant first order conditions, equations (27) and (28), respectively. That is, the left hand side of both equations are equal, thus right hands sides have to be equal as well. Organizing terms yields the relationship between entrant and incumbent per product line inventor employments, i.e. equation (1.16). Finally, we can derive equation (1.17) by using R&D production technologies (1.10) and (1.17).

D Evolution of firm size distribution

The mass of product lines owned by firms located in i is denoted by $\psi_i(t)$. To derive the system of equations that govern the time evolution of $\psi_i(t)$, we start with $\mathcal{P}_i(n,t)$ which denotes the measure of firms that are located in i and own n products. $\psi_i(t)$ and $\mathcal{P}_i(n,t)$ are related as follows

$$\psi_i(t) \equiv \sum_{n=1}^{\infty} n\mathcal{P}_i(n,t) \tag{29}$$

Unlike Klette and Kortum [2004], the firm size distribution has a spatial angle in the sense that we need to keep track of firm location which is the location where the firm locates its immobile R&D lab at the time of entry.

For $n = 2, 3, \ldots$, we have

$$\dot{\mathcal{P}}_i(n,t) = -n\big[x(t) + z_i(t)\big]\mathcal{P}_i(n,t) + (n-1)z_i(t)\mathcal{P}_i(n-1,t) + (n+1)x(t)\mathcal{P}_i(n+1,t) \tag{30}$$

Rate of change in $\mathcal{P}_i(n,t)$ equals three terms. First term represents outflows due to either creative destruction or incumbent firms' own innovation. The second term represents inflows due to innovations of firms with $n-1$ products. The third term also represents inflows to state n due to the fact that firms with $n+1$ products lose one product line as a result of creative destruction.

For state $n = 1$, we have the following equation

$$\dot{\mathcal{P}}_i(1,t) = -\big[x(t) + z_i(t)\big]\mathcal{P}_i(1,t) + \tilde{z}_i(t)\tilde{\psi}_i(t) + 2x(t)\mathcal{P}_i(2,t) \tag{31}$$

Similarly, first term represents outflows from the state. Second term stands for rate of entry that takes place in same location. Finally, third term represents inflows from state $n = 2$ due to creative destruction.

While the total mass of product lines is normalized to one, the total mass of firms is endogenous and evolves over time. Denoting by $M_i(t)$ the mass of firms located in i at time t, we can derive the following equation

$$\dot{M}_i(t) = -x(t)\mathcal{P}_i(1,t) + \tilde{z}_i(t)\tilde{\psi}_i(t) \tag{32}$$

Rate of change in $M_i(t)$ equals to inflow due to entry in i, the second term, minus outflow due to creative destruction. Note that firms losing their last product line exits the economy.

E Equilibrium production worker wage rate

Demand equation for intermediate good ν from final good producers in a location i is as follows

$$p(\nu,t)k_i(\nu,t) = (1-\beta)Y_i(t)$$

Replacing optimal pricing rule $p(\nu,t) = \lambda a(\nu,t)^{-1}$ yields $k_i(\nu,t) = \lambda^{-1}a(\nu,t)(1-\beta)Y_i(t)$. Replacing this into final good production function and using the constant $\bar{\mathcal{A}} = \left(\frac{\lambda}{1-\beta}\right)^{1-\beta}$, we can solve for the output in i as

$$Y_i(t) = \mathcal{A}(t)^{\frac{1-\beta}{\beta}} L_i(t)$$

where $\mathcal{A}(t) = \exp\left(\int_0^1 \log a(\nu,t)d\nu\right)$ is aggregate productivity index of the economy. Then total output equals

$$Y(t) \equiv \sum_i Y_i(t) = \mathcal{A}(t)^{\frac{1-\beta}{\beta}} \bar{L}$$

Demand for labor is given by $W_i^L(t)L_j(t) = \beta Y_i(t)$. Replacing output and solving for wage results in

$$W_i^L(t) = \beta \mathcal{A}(t)^{\frac{1-\beta}{\beta}}$$

Substituting $\mathcal{A}(t)^{\frac{1-\beta}{\beta}}$ with $\frac{Y(t)}{\bar{L}}$ yields equation (1.21) in the main text.

F *Growth rate of* $\mathcal{A}(t)$

From definition of $\mathcal{A}(t)$, we have

$$\log \mathcal{A}(t) = \int_0^1 \log a\,(\nu, t)\, d\nu$$

Then

$$\frac{\dot{\mathcal{A}}(t)}{\mathcal{A}(t)} = \frac{d \log \mathcal{A}(t)}{dt} = \lim_{dt \to 0} \frac{\log \mathcal{A}(t+dt) - \log \mathcal{A}(t)}{dt}$$

Using definition, we can show

$$\log \mathcal{A}(t+dt) - \log \mathcal{A}(t) = \int_0^1 \left[\log a\,(\nu, t+dt) - \log a\,(\nu, t)\right] d\nu$$

Productivity of a product line ν increases by a proportionality factor of $\lambda > 1$ as a result of creative destruction. The equilibrium rate of creative destruction is $x(t)$. Therefore, for a small time interval of dt, we can write

$$a\,(\nu, t+dt) = \begin{cases} \lambda a(\nu, t) & \text{with probability } x(t)dt \\ a(\nu, t) & \text{with probability } 1 - x(t)dt \end{cases}$$

Thus, $\log a(\nu, t+dt) - \log a(\nu, t)$ is also a random variable with the following

$$\log a(\nu, t+dt) - \log a(\nu, t) = \begin{cases} \log \lambda & \text{with probability } x(t)dt \\ 0 & \text{with probability } 1 - x(t)dt \end{cases}$$

Integrating over all product lines $\nu \in [0, 1]$, we have

$$\frac{\dot{\mathcal{A}}(t)}{\mathcal{A}(t)} = \lim_{dt \to 0} \frac{\log(\lambda)\, x(t)dt}{dt} = \log(\lambda)\, x(t)$$

as given in the main text.

G Allocation of profits across agents

Let $D(t)$ denote the total profits paid to agents. It equals to sum of profits of intermediate good firms after subtracting R&D costs (including entrants). Then

$$D(t) = \left(1 - \lambda^{-1}\right)(1 - \beta)\, Y(t) - \sum_i W_i^R(t)(1 - s_i)R_i(t)$$

From the distribution of profits, we also have

$$D(t) = \sum_i d(t) W_i^L(t) L_i(t) + \sum_i d(t) W_i^R(t) R_i(t)$$

Then solving for $d(t)$ yields

$$d(t) = \frac{\left(1 - \lambda^{-1}\right)(1 - \beta)\, Y(t) - \sum_i W_i^R(t)(1 - s_i)R_i(t)}{\sum_i W_i^L(t) L_i(t) + \sum_i W_i^R(t) R_i(t)} \tag{33}$$

H Proof of Theorem 1.3.3

I prove this theorem under the conjecture that $z_i(t) = z$, $\tilde{z}_i = \tilde{z}$ for all i, and $x(t) = x$ as assumed in the main text. If r_i is constant over time in BGP equilibrium as conjectured, then $\tilde{r}_i$ is also constant satisfying $\tilde{r}_i = \frac{r_i}{F}$ (Theorem 1.3.2). Total inventor employment in a location i equals the sum of inventors in incumbent and entrant firms in that location. Under the conjecture that ψ_i and $\tilde{\psi}_i$ are constant over time in BGP, then $R_i = \psi_i r_i + \tilde{\psi}_i \tilde{r}_i$ is also constant, and

$$\begin{aligned} x &= \sum_i \psi_i z_i + \sum_i \tilde{\psi}_i \tilde{z}_i \\ &= z \sum_i \psi_i + \tilde{z} \sum_i \tilde{\psi}_i \\ &= z + \tilde{z} \end{aligned} \tag{34}$$

As $\sum_i \psi_i = \sum_i \tilde{\psi}_i = 1$. System of equations (30), (31) and (32) admit a stationary solution $\mathcal{P}_i(n,t) = \mathcal{P}_i(n)$ such that[20]

$$\mathcal{P}_i(n) = \frac{\tilde{\psi}_i \tilde{z} z^{n-1}}{n x^n}, \qquad n = 1, 2, \ldots \tag{35}$$

Then

$$\begin{aligned} \psi_i &= \sum_{n=1}^{\infty} \mathcal{P}_i(n) n = \sum_{n=1}^{\infty} \frac{\tilde{\psi}_i \tilde{z}_i z_i^{n-1}}{n x^n} n = \tilde{\psi}_i \frac{\tilde{z}}{z} \sum_{n=1}^{\infty} \left(\frac{z}{x}\right)^n \\ &= \tilde{\psi}_i \frac{\tilde{z}}{z} \left(\frac{1}{1 - \frac{z}{x}} - 1 \right) \\ &= \tilde{\psi}_i \end{aligned} \tag{36}$$

Second line follows from (34) given that $z > 0$ and $\tilde{z} > 0$ in BGP equilibrium.

20. See Klette and Kortum [2004] for details.

Inventor market clearing implies

$$\begin{aligned} R_i &= \psi_i r_i + \tilde{\psi}_i \tilde{r}_i \\ &= \left(1 + F^{-1}\right) \psi_i r_i \\ &= \frac{1+F}{F} \psi_i \frac{z^\theta}{\alpha_i} \end{aligned}$$

which implies

$$\psi_i = \frac{F}{1+F} z^{-\theta} \alpha_i R_i \tag{37}$$

We can solve for z using $\sum_i \psi_i = 1$. That is,

$$\sum_i \psi_i = 1 = \frac{F}{1+F} z^{-\theta} \sum_i \alpha_i R_i$$

which gives

$$z = \left[\frac{F}{1+F} \sum_i \alpha_i R_i \right]^{\frac{1}{\theta}} \tag{38}$$

Using (38) and (37), we can show that

$$\psi_i = \tilde{\psi}_i = \frac{\alpha_i R_i}{\sum_i \alpha_i R_i} \tag{39}$$

as stated in Theorem 1.3.3, satisfying the initial conjecture that ψ_i and $\tilde{\psi}_i$ are time invariant.

I Proof of Theorem 1.3.4

From (1.26) and $z_i = z$ for all i, it follows that

$$v_i = v \equiv \frac{\pi}{\rho - g + x - \frac{\theta - 1}{\theta} z}$$

First order condition of incumbent firm maximization problem (27) and the fact that $\mathcal{V}_i(n,t) = nvY(t)$ in BGP imply that

$$W_i^R(t)(1-s_i) = \alpha_i^{\frac{1}{\theta}}\frac{1}{\theta}r_i^{\frac{1}{\theta}-1}vY(t)$$

in BGP. Defining normalized inventor wage $w_i^R(t) = \frac{W_i^R(t)}{Y(t)}$, and as $z = z_i = (\alpha_i r_i)^{\frac{1}{\theta}}$ for all i, we can rewrite above equation as

$$w_i^R(t)(1-s_i) = z^{1-\theta}\frac{1}{\theta}\alpha_i v$$

for all t. It immediately follows that normalized inventor wage is constant in BGP and equals to

$$w_i^R = \frac{1}{\theta}z^{1-\theta}v\frac{\alpha_i}{1-s_i}$$

as stated in Theorem 1.3.4.

J Proof of Theorem 1.3.5

In BGP, all consumption rates grow with $g = \frac{1-\beta}{\beta}\log(\lambda)x$. Thus, relative consumptions across locations are time invariant. Let $C_{ij}^T \equiv \frac{C_i^T(t)}{C_j^T(t)}$ for type $T = W, R$. Define normalized value functions $u_i^T(\varepsilon,t) \equiv \frac{\mathcal{U}_i^T(\varepsilon,t)}{C_i^T(t)}$ and $\bar{u}_i^T(t) \equiv \frac{\bar{\mathcal{U}}^T(t)}{C_i^T(t)}$. First we can show that

$$\bar{u}_i^T(t) = \int \max_j \left(u_j^T(e_j,t)C_{ji}^T\right) f_\epsilon(\mathbf{e})d\mathbf{e} \tag{40}$$

Organizing (1.6), we can show that normalized value function $u_i^T(\varepsilon,t)$ satisfies the following functional equation

$$[\rho+\zeta-g]u_i^T(\varepsilon,t) = A_i\varepsilon + \zeta\bar{u}_i^T(t) + \partial_t u_i^T(\varepsilon,t) \tag{41}$$

Following time invariant solutions to normalized agent value functions satisfy the system of equations given by (40) and (41)

$$u_i^T(\varepsilon) = \frac{A_i\varepsilon + \zeta\bar{u}_i^T}{\rho + \zeta - g} \quad \text{and} \quad \bar{u}_i^T = \int \max_j \left(u_j^T(e_j)C_{ji}^T\right) f_\epsilon(\mathbf{e})d\mathbf{e}$$

We can recover agent value functions simply using definitions

$$\mathcal{U}_i^T(\varepsilon, t) = \frac{A_i\varepsilon C_i^T(t) + \zeta\bar{\mathcal{U}}^T(t)}{\rho + \zeta - g} \tag{42}$$

Replacing value function (42) into the migration problem (1.4) yields

$$\max_j \mathcal{U}_j^T(e_j, t) = \frac{\max_j \left(A_j e_j C_j^T(t)\right) + \zeta\bar{\mathcal{U}}^T(t)}{\rho + \zeta - g} \tag{43}$$

Substituting (43) into the definition of $\bar{\mathcal{U}}^T(t)$, equation (1.5), results in

$$\bar{\mathcal{U}}^T(t) = \frac{1}{\rho + \zeta - g}\left(\mathbb{E}_\epsilon\left[\max_j \left(A_j e_j C_j^T(t)\right)\right] + \zeta\bar{\mathcal{U}}^T(t)\right) \tag{44}$$

In order to take expectation in (44), we need to derive the distribution of the maximum term. The assumption that locations tastes are independently distributed Frechet, i.e. $\varepsilon_i \sim Frechet(\xi, 1)$, implies

$$\max_j \left(A_j e_j C_j^T(t)\right) \sim Frechet\left(\xi, \left[\sum_{j=1}^{N}\left(A_j C_j^T(t)\right)^\xi\right]^{\frac{1}{\xi}}\right) \tag{45}$$

Taking expectation in (44) and solving for $\bar{U}^T(t)$ yields

$$\bar{\mathcal{U}}^T(t) = \frac{1}{\rho - g}\Gamma\left(1 - \frac{1}{\xi}\right)\left[\sum_{j=1}^{N}\left(A_j C_j^T(t)\right)^\xi\right]^{\frac{1}{\xi}} \tag{46}$$

Finally substituting $\bar{U}^T(t)$ into the agent value function in (42) results in

$$\mathcal{U}_i^T(\varepsilon,t) = \frac{A_i\varepsilon C_i^T(t)}{\rho+\zeta-g} + \frac{\zeta}{(\rho+\zeta-g)(\rho-g)}\Gamma\left(1-\frac{1}{\xi}\right)\left[\sum_{j=1}^{N}\left(A_jC_j^T(t)\right)^{\xi}\right]^{\frac{1}{\xi}}$$

as stated in Theorem 1.3.5.

K Proof of Theorem 1.3.6

Provided the analytical solution for the agent value function from Theorem 1.3.5, the migration choice given by (1.4) simplifies such that

$$i^T(t)^\star = \arg\max_j \mathcal{U}_j^T(e_j,t) = \arg\max_j \left(A_je_jC_j^T(t)\right) \tag{47}$$

as the second term on the right hand side of equation (1.29) is independent of locations. Moreover, consumption is proportional to wage rate, i.e.

$$C_j^T(t) = [1+d-\tau]\,w_i^T Y(t) \tag{48}$$

where d is the proportionality factor of profits allocated to agents to wages, given by (33), and τ is the labor income tax rate, given by (1.25).[21] As d and τ are independent of locations,

21. In BGP, $d(t)$ is time invariant. In equation (33), $Y(t)$ cancels from both numerator and denominator of the expression. As a result, in BGP, we have

$$d = \frac{\left(1-\lambda^{-1}\right)(1-\beta) - \sum_i w_i^R(1-s_i)R_i}{\sum_i w_i^L L_i + \sum_i w_i^R R_i}$$

after a conjecture that worker and researcher populations are stable in BGP. This conjecture holds in BGP as a result of the migration decision given by (47) as proven below. Moreover, given same conjecture, we can show that $\tau(t)$ is independent of time in BGP. That is, following from equation (1.25),

$$\tau = \frac{\sum_{i=1}^N s_i w_i^R R_i}{\sum_{i=1}^N w_i^L L_i + w_i^R R_i}$$

replacing (48) into the migration problem (47) delivers equation (1.30).

In BGP, an agent from i, who has just drawn taste shocks $\{e_m\}_{m=1}^N$, migrates to j if and only if

$$A_j e_j w_j^T \geq A_m e_m C_m^T, \qquad \forall m = 1, \ldots, N$$

Let γ_{ji}^T denote the share of agents of type T in i moving to j among who draw taste shocks. Then,

$$\begin{aligned}
\gamma_{ji}^T &= \mathbb{P}\left\{A_j e_j w_j^T \geq A_m e_m w_m^T,\ \forall m \neq j\right\} \\
&= \frac{\left(A_j w_j^T\right)^\xi}{\sum_{m=1}^N \left(A_m w_m^T\right)^\xi}, \qquad \forall i = 1, \ldots, N
\end{aligned}$$

The second line follows from the property that taste shocks are distributed Frechet. Notice that γ_{ji}^T does not vary with i. Hence we can assert that $\gamma_{ji}^T = \gamma_j^T$ for all i.

Let M_j^T denote the mass of type-T agents located in j.[22] Then law of motion of this variable depends on migration flows γ_j^T such that

$$\begin{aligned}
\dot{M}_j^T &= \underbrace{\sum_{i=1}^N \gamma_j^T \zeta M_i^T}_{\text{Inflow}} - \underbrace{\zeta M_j^k}_{\text{Outflow}} \\
&= \zeta\left(\gamma_j^T M^T - M_j^T\right)
\end{aligned}$$

where M^T is the total population of type-T, i.e. $M^T = \sum_j M_j^T$. As $\dot{M}_j^T = 0$ in BGP, we have $M_j^T = \gamma_j^T M^T$. That is,

$$\begin{aligned}
L_j &= \gamma_j^L \bar{L} \\
R_j &= \gamma_j^R \bar{R}
\end{aligned}$$

22. $M_j^L = L_j$ and $M_j^R = R_j$.

Therefore, the initial conjecture that population shares are time invariant holds true.

Finally, we can replace the equilibrium wage rates into the expression for γ_j^T. Firstly, $w_j^L = \frac{\beta}{L}$ for all j. Hence,

$$\gamma_j^L = \frac{A_j^{\xi}}{\sum_{m=1}^{N} A_m^{\xi}}$$

as given by (1.31). Using equation (1.3.4), we can also prove that

$$\gamma_j^R = \frac{A_j^{\xi}\left(\frac{\alpha_j}{1-s_j}\right)^{\xi}}{\sum_{m=1}^{N} A_m^{\xi}\left(\frac{\alpha_m}{1-s_m}\right)^{\xi}}$$

Combining the expressions for γ_j^L and γ_j^R proves equation (1.32) given by Theorem 1.3.6.

L Proof of Theorem 1.3.7

The expression for z in BGP given by (1.34) is derived in Section H (equation (38)).

Total rate of innovation generated in a location i, x_i, equals to the sum of two terms, first of which is the product of mass of product lines owned by incumbent firms located in i and rate of innovation per product line z. The second term is the product of the mass of entrants located in i and the rate of innovation per entrant, i.e. $\tilde{z}$. Thus,

$$x_i = \psi_i z + \tilde{\psi}_i \tilde{z}$$

In Section H equation (37), it is shown that $\psi_i = \tilde{\psi}_i = \frac{F}{1+F} z^{-\theta}\alpha_i R_i$. Moreover, we know $\tilde{z} = z/F$. Thus

$$\begin{aligned} x_i &= \left(1 + F^{-1}\right)\psi_i z \\ &= \frac{1+F}{F}\frac{F}{1+F} z^{-\theta}\alpha_i R_i z \\ &= z^{1-\theta}\alpha_i R_i \end{aligned}$$

as stated in Theorem 1.3.7 equation (1.35).

As shown by equation (34), $x = z + \tilde{z} = \frac{1+F}{F} z$. Finally, equation (1.37) follows from equations (1.23) and (1.36).