Faculty of Business and Law

Working Papers in Economics & Finance
No. 2024-07

Innovation and Regional Development: The Impact of Patenting on Labor Market Outcomes

Ali Sina Önder, University of Portsmouth

Sascha Schweitzer, Reutlingen University

Olga Tcaci, TUD Dresden University of Technology

Portsmouth Business School
https://www.port.ac.uk/about-us/structure-and-governance/organisational-structure/our-academic-structure/faculty-of-business-and-law/portsmouth-business-school

Innovation and Regional Development: The Impact of Patenting on Labor Market Outcomes

Ali Sina Önder* Sascha Schweitzer† Olga Tcaci‡

5 November 2024

Abstract

We estimate the impact of technological innovation on regional labor market outcomes. Our identification strategy exploits pre-reunification complementarities in innovation between East and West Germany. We employ individual-level data from the German Socio-Economic Panel to analyze labor market outcomes. Individuals' income in West German counties with pre-reunification complementarities increased by 1.3%-1.5% on average after reunification. The effect is amplified when disentangling for different occupations: Income increases by 27%-29%, self-employment increases significantly, unemployment remains unaffected. The use of East German know-how in West German patents after reunification is driven by the migration of East German inventors to West German counties.

JEL classification: J24; O31; O33; R11.

Keywords: Economic Development, Patent Analysis, Knowledge Complementarities, Occupations, German Reunification.

*Faculty of Business and Law, University of Portsmouth, Portsmouth PO1 3DE, UK. ali.onder@port.ac.uk

†ESB Business School, Reutlingen University, 72762 Reutlingen, Germany. sascha.schweitzer@reutlingen-university.de

‡Faculty of Business and Economics, TUD Dresden University of Technology, 01062 Dresden, Germany. olga.tcaci@tu-dresden.de

1 Introduction

There is a general consensus that innovation is a powerful engine of economic development. Extensive literature on scientific progress (Bilbao-Osorio and Rodríguez-Pose, 2004; Rodríguez-Pose and Crescenzi, 2008), diffusion of knowledge and ideas (B. F. Jones, 2009; C. I. Jones, 2023; Romer, 2009; Weitzman, 1998) and technological advancements measured by patents (Akçomak and ter Weel, 2009; Akcigit et al., 2017; Bloom et al., 2020) documents the contribution of innovation for regional economic performance. More recent evidence attributes rising income inequality to the "innovation-led growth" (Aghion et al., 2019; Breau et al., 2014; C. I. Jones and Kim, 2018). This, in turn, leads to increasing disparities within and across regions (Iammarino et al., 2019). We investigate the causal effects of regional innovation on individuals' employment patterns across and within regions, focusing on the differential effects of innovation across occupations.

Establishing causality in this context is challenging mainly due to reverse causality in the positive association between regional innovation and growth. We address this shortcoming by deriving the causal link between regional innovation and labor market outcomes of individuals in the context of the German separation and sudden reunification more than 40 years later. This historical event provides a compelling framework to isolate the causal effects of regional innovation on real economic outcomes. The unexpected fall of the Berlin Wall in 1989 and the subsequent reunification of the two German states in 1990 involved a considerable restructuring of East Germany (Sabel, 1993). As a consequence of this exogenous shock, a massive inventor migration from East to West Germany occurred (Agrawal et al., 2016; Borjas and Doran, 2012; Ferrucci, 2020; Hnermund and Hipp, 2024). This allows us to exploit the variation in the ability of West German counties to make use of the unexpectedly available technological know-how from East Germany.

We quantify regional innovation activities using regional patent applications, which serve as a proxy for the technological knowledge content of innovative activities (Griliches, 1990; Hall et al., 2001), capturing creative and application-based innovation patterns (Capello and Lenzi, 2013, 2019). We use the European Patent Office's (EPO) Worldwide Patent Statistical Database (PATSTAT) patent data from 1965 to 2004.[1] From the PATSTAT information on the location of inventors and patent application authorities, we identify patents registered and inventors based[2] in East and West Germany, as well as collaborative patents between East and West German scientists. Additionally, we retain information on the technological content of the patent according to the International Patent Classification (IPC) taxonomy.

Our identification strategy exploits complementarities between the East and West German knowledge found in their collaborative patents before the reunification. Collaborations between East and West German inventors that led to patent applications were limited before reunification due to administrative hurdles between authoritarian East Germany and democratic West Germany. Despite the great costs associated with the collaborative activities between the two countries, their existence indicates that such efforts must be worthwhile and highly valuable.[3] To operationalize this information, we use the fractional count method as in Ferrucci (2020) and identify the top 75% of IPCs of pre-reunification East, West, and East-West German patent applications. We then identify unique East German IPCs that appear along with unique West German IPCs on the collaborative patents. We define these IPCs as complementary IPCs because they embody West German know-how that has been successfully complemented by East German know-how. In our empirical setting, regions with pre-reunification knowledge complementarities possess a comparative advantage in attracting and utilizing East German know-how post-reunification. Considering the previously discussed constraints in cooperation between the two countries and the unforeseen event of reunification, the presence of complementary IPCs in West German counties can be regarded as exogenous. We compute the share of complementary IPCs from the total IPCs at the county level. A county's patent portfolio during the 1980s represents the treatment of our analysis. Using a difference-in-differences framework for 1985-1989 and 1992-1996, we estimate whether a region with pre-existing knowledge complementarity registers positive externalities for the labor market outcomes of individuals living in the respective region. To address this, we use the population-representative individual-level data from the German Socio-Economic Panel (SOEP) (Goebel et al., 2019) containing information on the socio-economic aspects such as labor income, unemployment and self-employment for residents of West German counties starting from 1984. To the best of our knowledge, this is the first analysis combining SOEP data with the comprehensive PATSTAT data.

Our results show that West German counties with strong pre-reunification complementarities to East German knowledge have a positive and significant increase in their use of East German know-how in their patents after the reunification. Next, we provide evidence that this effect is primarily driven by the relocation of East German inventors into these counties. This confirms that our identification strategy correctly captures changes in patenting activities that result from collaborations with East German inventors. This also ensures that we rule out reverse causality so that our findings on employment patterns of individuals are the result of regional patenting activities rather than their cause.

In the next step of our analysis, we show that individuals labor income in West German counties with strong complementarities to the East increased significantly. A one standard deviation increase in a county's pre-reunification complementarity to East German knowledge leads to an increase in its residents' income by 1.3% to 1.5% on average. We estimate a positive but noisy difference-in-differences effect on individuals self-employment probability, and we find zero effect on the duration of unemployment. This empirical analysis raises concerns about the county-level average effects. Namely, the effect of innovation cannot be homogeneous across all occupations in our sample (Breau et al., 2014). To address this concern, we create an annual proximity measure between each individuals occupation and the countys patent portfolio based on IPCs of patents. The baseline ordinary least squares results show a positive and statistically significant association between this proximity measure and income. This suggests that individuals employed in sectors where a higher share of patents are registered tend to earn higher incomes after reunification. Thereafter, we instrument the proximity measure with the pre-existing knowledge complementarities. In a two-stage least squares setting, we regress individuals income, unemployment, and self-employment on the instrumented proximity measure. The results reveal a statistically and economically significant effect on income: A one standard deviation increase in the instrumented proximity increases income by 27% to 29%. The probability of being self-employed increases significantly, whereas the duration of unemployment remains unaffected. In our robustness analysis, we confirm that the knowledge complementarities between East and West Germany exclusively drive our results. To further limit the probability of the potential reverse causality, all our explanatory variables are lagged by up to three years. We also control for the possibility of pre-existing East German knowledge in West German counties. In addition, we validate the robustness of our results using an alternative level of analysis, specifically the labor market areas. In comparison to previously presented county-level analysis, labor market areas also account for commuting zones in Germany and further economic aspects (Kropp and Schwengler, 2011).

This study contributes to several strands of the literature on the economic returns to innovation. First, we contribute to the well-established endogenous growth literature (Aghion et al., 2019; Aghion and Howitt, 1992; Akcigit et al., 2017; Romer, 1990) providing novel insights on (the value of) patent publications (Hegde et al., 2023; Sampat and Williams, 2019), considering the patent quality (De Rassenfosse et al., 2014; Harhoff et al., 1999) for innovative productivity and regional economic development (Capello and Lenzi, 2019; Del Monte et al., 2020; Griliches, 1990).

Second, our work adds to the literature analyzing the collaborations between inventors and the resulting benefits. While previous literature focuses mainly on the team sizes (Agrawal et al., 2016), the prominence of the collaborators (Azoulay et al., 2019), and the migration of fellow scientists (Borjas and Doran, 2015; Ferrucci, 2020; Moser et al., 2014), we identify and analyze the complementarities in the knowledge space of collaborations. This allows us to localize the effects of collaboration, directly connecting to the third aspect of our contribution.

Third, our results have implications for the increasing literature on exposure of occupations to innovational progress (Autor and Dorn, 2013; Brynjolfsson and McAfee, 2014; Kogan et al., 2021, 2017; Mann and Püttmann, 2023; Mokyr et al., 2015; Webb, 2019). Existing literature focuses on a specific type of innovational or technological progress, such as computerization (Frey and Osborne, 2017), machine learning (Brynjolfsson et al., 2018), and, more recently, artificial intelligence (Acemoglu, 2021). This article considers different types of innovational progress and attributes their effects to a particular occupation category.

The remainder of the paper is structured as follows. Section 2 outlines our conceptual framework and provides the historical context for our analysis. Section 3 describes the data preparation and analysis. In section 4, we provide a detailed explanation of our identification strategy. Section 5 summarizes our main results. Section 6 presents robustness analysis. Section 7 concludes.

2 Conceptual Framework and Historical Background

Innovation and Economic Development. The underlying fundamental mechanism of the relationship between innovation, economic growth, and productivity has been widely analyzed in the literature. The main channel on how innovation can promote growth has been argued to be that research and development (R&D) investment and human capital accumulation interact to continuously raise productivity and thus generate economic growth (Aghion and Howitt, 1992; Romer, 1990). This process requires adequate public policies (Bloom et al., 2019; Hall and van Reenen, 2000; Jaffe and Lerner, 2011) and institutions (Aghion et al., 2015; North, 1990) nationwide.

Understanding the economic impact of regional innovation activities is crucial in the context of increasing regional economic disparities. Territorial innovation models contribute in this sense by providing important insights into the requirements that enable regional innovation (Asheim, 2012; Capello and Lenzi, 2013; Fritsch, 2001), emphasizing the local nature of innovation. Previous studies in regional economics and economic geography delivered evidence of the positive association of regional income with regional innovation activities (Capello and Lenzi, 2019; Crescenzi and Rodríguez-Pose, 2011). Since spillover effects are geographically highly localized (Audretsch and Feldman, 1996; Feldman and Audretsch, 1999; Jaffe et al., 1993), local innovation activities are vital for regional economic outcomes (Boschma and Fritsch, 2009). For instance, Bottazzi and Peri (2003) show that a doubling of R&D spending in a region almost doubles patent applications in that region, while it increases patent applications only by 3% in surrounding regions. Moreover, the effects of innovation activities are not necessarily homogeneous across income levels (Aghion et al., 2019; Hémous and Olsen, 2022), which is mainly due to the heterogeneous effects of innovation on labor productivity (Kogan et al., 2021, 2017; Rocchetta et al., 2022) and occupational structure (Acemoglu and Restrepo, 2020; Autor and Dorn, 2013; Mann and Püttmann, 2023).

Griliches (1990) argues that patents serve as a tangible and quantifiable measure of innovation, representing legally recognized inventions and providing a quantitative indicator of technological progress. In addition, patent data provide rich, standardized information consistently recorded across different jurisdictions and over time, thus allowing for comparative analysis across regions, as noted by Jaffe and Trajtenberg (2002). As a measure of knowledge production, patent data is also a reliable measure of income growth representative of a region (Acs et al., 2002). While R&D spending is another indicator of innovation, it often lacks the specificity and outcome-oriented focus of patent data. Following this reasoning, we use patent data in our study to accurately capture regional innovation activity.

An area of growing interest, which we address in our study from a regional economic perspective, is the relationship between patenting activities and occupational structures. The impact of patenting on job creation has been widely studied, particularly in high-tech and knowledge-intensive sectors. Bessen (2015) shows that patenting firms employ more in R&D, engineering, and other technical occupations because firms patenting activities signal firm growth and attract investment, leading to expansion of high-skill occupations. Thus, patenting activities may lead to a shift in occupational demand towards high-skilled occupations. The resulting shift in wages is not the only source of occupational heterogeneity. Hunt and Gauthier-Loiselle (2010) suggest that patenting can contribute to wage dispersion within firms and industries, as those with the skills to contribute to or manage intellectual property are rewarded more generously. Patent-intensive firms often pay higher wages, particularly for skilled workers involved in the innovation process. Therefore, regions with higher patenting activity will experience greater employment growth in high-skilled occupations (Audretsch and Feldman, 1996). For example, Breau et al. (2014) find a positive correlation between innovative activity and wage inequality within Canadian cities.

This study complements previous regional economics literature by establishing causal mechanisms between regional patenting activities and individuals income, employment, and occupations. We analyze the overlap between individuals' occupational sectors and their region's patenting sectors to identify the sector exposure of occupations to innovation activities. Our data allow us to follow individuals over time and provide informative results on the implication of the proximity between innovation and occupation on the individuals' income.

Science and Innovation in East vs. West Germany. The empirical analysis relies on regional differences in innovation, which allows us to establish causality by exploiting the natural experiment of German reunification. To corroborate the validity of our analysis, this part provides historical background on the science and innovation system in East and West Germany before German reunification.

Following the Second World War, Germany was divided into West Germany (Federal Republic of Germany) that was built upon the democratic principles of the free market, and East Germany (German Democratic Republic) that was governed by an authoritarian regime implemented by the Soviet Union with a rigid and centralized state-owned economy. This division lasted from 1949 until the reunification in 1990. The four decades of division between the two countries have had a notable impact on both countries. The consequences of the divide continue to perpetuate contemporary political preferences (Avdeenko, 2018; Weisskircher, 2020), economic (Beblo and Görges, 2018; Laudenbach et al., 2020; Wyrwich, 2019) and socio-cultural (Brosig-Koch et al., 2011; Heineck and Süssmuth, 2013; Möhlmann, 2014; Rainer and Siedler, 2009) disparities.

The highly rigid innovation system in East Germany was primarily determined by a combination of political interests and industrial needs in the context of the Cold War with a focus on fundamental research areas such as mathematics and physics (Agrawal et al., 2016; Chan et al., 2022; Ferrucci, 2020; Sabel, 1993). Like university science, the industrial sector research efforts were discouraged mainly by the lack of funding and materials shortages (Sabel, 1993). Despite the inefficiencies in the innovation system of East Germany compared to West Germany (Fritsch et al., 2023), and peculiarities such as the introduction of the "economic patent" (*Wirtschaftspatent*) in 1950 (Glitz and Meyersson, 2020; Hnermund and Hipp, 2024), East German patent applications followed the common international patenting rules Hoisl et al. (2016), as we also show in Figures A3 and A4 in the Appendix[4]. Moreover, Glitz and Meyersson (2020) show that both overt means, such as "exclusive patents" (*Ausschlieungspatent*), and covert methods, such as espionage, were used by East Germany to target West German knowledge. Meanwhile, West Germany experienced the famous "Wirtschaftswunder" (economic miracle), benefiting from the democratic free market economy that also defined the more productive innovation system based on the Humboldtian academic system (Fritsch et al., 2023; Fritsch and Wyrwich, 2021).

Despite the restrictions on scientific collaborations imposed during the early years of the Cold War, a scientific exchange was facilitated between the two countries for the first time by the treaty *Grundlagenvertrag* in 1972. Later, with the Conference on Security and Cooperation in Europe (CSCE) treaty of 1975, the Cultural Exchange Agreement of 1985, and the agreement of scientific and technological cooperation (WTZ agreement) of 1987 in the following years, substantial cooperation became possible (Önder, 2022; Sabel, 1993).

Due to geographic proximity, West Germany was the primary beneficiary of knowledge transfer from East Germany after the reunification. General similarities such as the same mother tongue, cultural heritage, and high education levels made the human capital transferable between East and West Germany (Prantl and

Spitz-Oener, 2020).[5] Therefore, we focus on analyzing the impact of West Germany's opportunity to harness this large pool of knowledge from East Germany after reunification, particularly in expanding the knowledge frontier and regional economic development.

3 Data

3.1 Innovation: PATSTAT

We use a subset of the 2022 version of the PATSTAT dataset. In contrast to several studies mentioned above, we do not restrict our analysis to one patent authority. The main inclusion criterion for patent applications is the presence of an East or West German inventor with at least two patents between 1965 and 2004. To avoid duplicates, we restrict the analysis to priority filings. Furthermore, we exclude utility models, PCT applications, provisional applications, design patents, plant patents, and artificial applications. From the original PATSTAT dataset, we derive additional variables such as a custom person ID for inventors based on matching their names, their earliest known city of residence based on these IDs, their gender based on first names, and patent renewals based on legal events. For comparability across patent authorities, we normalize the number of renewals for each patent by the maximum number of renewals within a year and the same patent authority.[6]

We link patents to specific counties (NUTS3-level) using the inventors' addresses, focusing on East and West German patent offices. Because of this regional focus, we did not use the OECD REGPAT Database, which is limited to EPO, PCT, and USPTO patents (Maraut et al., 2008). Our dataset includes NUTS3-level information from 1965 to 2004 provided by PATSTAT, enriched with postal code data from patent applications. Before the German reunification, two overlapping postal systems existed in East and West Germany. This requires manual checks to resolve ambiguities. After German reunification, a unified five-digit postal system was established. We use data from https://www.alte-postleitzahlen.de (accessed August 22, 2023) to map old to new postal codes. We convert NUTS3-level data to county codes using official information and complement it with county codes derived from new postal codes using OpenStreetMap data (accessed August 22, 2023), thus enhancing the datasets accuracy and utility.

From the available detailed information, we identify the East-West German collaborations using the address information, where at least one inventor is based in an East German county and the other co-inventor(s) in West Germany. In addition to the authors' information, patents contain information on the patent-specific technological classification,[7] size of inventor teams, and further indices of patent quality such as forward citations of patents (De Rassenfosse et al., 2014; Moser et al., 2018) and patent renewals.

Our final complete patent dataset results in over one million distinct patent applications and over 2.5 million distinct observations (considering distinct patent applications and distinct inventor locations) for East and West German registered patents from 1965 to 2004.[8] The reader is invited to refer to Appendix C for further details on our inclusion criteria, the regionalization, and variables derived from PATSTAT.

3.2 Economic Performance: German Socio-Economic Panel (SOEP)

We use individual data from SOEP for our analysis of labor market outcomes. The annual panel survey starting in 1984[9] for West German residents provides representative data for the German population (Goebel et al., 2019). Data is available at the individual and household level, including information on occupation, earnings, employment, and other socio-economic indicators. We construct a sample of West German respondents and use the additional restriction that their county of residence did not change after the reunification. In this way, we argue that the effect we are analyzing is not measured at the individual level but at the county level. Given the different regional reorganizations, SOEP converts the former West German county codes to the currently available version of the county code, which allows us to match our patent data.

For our analysis of regional economic development, we use information on current labor income and the number of months of official unemployment in the previous year, and we restrict the analysis to working-age individuals. Following Lichter et al. (2021), who analyze the variation in the labor market outcomes after the reunification using the discontinuities in surveillance intensity across county borders of the pre-reunification East Germany, we account for inflation by calculating real income in 2000 prices using the West German Consumer Price Index (CPI). We drop the bottom and top 1% of the income distribution. Similarly, our variable of unemployment duration is calculated for each individual as the ratio of months of unemployment over the total number of months in one period.

For working-age individuals, we also consider their probability of being self-employed. Such individuals are identified in the SOEP as self-employed without employees or self-employed with employees (SOEP, 2020).

3.3 Sectors of Occupations and Patents: NACE Rev. 1.1 and IPCs

SOEP reports individuals' occupations using a specific industry taxonomy, which is known as the nomenclature of economic activities (NACE). NACE is a statistical classification system used by the European Community to categorize economic activities. SOEP provides data for the NACE Rev. 1.1. As patents' technology classifications (IPCs) do not align directly with NACE classification, we use the concordance table provided by EUROSTAT (2001) and attribute to the IPCs classes of patents the corresponding NACE Rev. 1.1 industrial occupation classes. This allows us to link patent activity to the corresponding industrial occupations of individuals in the SOEP dataset and subsequently analyze the relationship between these variables.

The correspondence between the NACE and IPCs categories allows us to examine how individuals' occupation sectors align or mismatch with sectors of patenting activity in a given county in a given year. We refer to the sectors in a county that closely align with patenting activities in that county's local innovation sectors.[10] We provide a detailed explanation of the composition of our SOEP sample and the main variables in Appendix B.

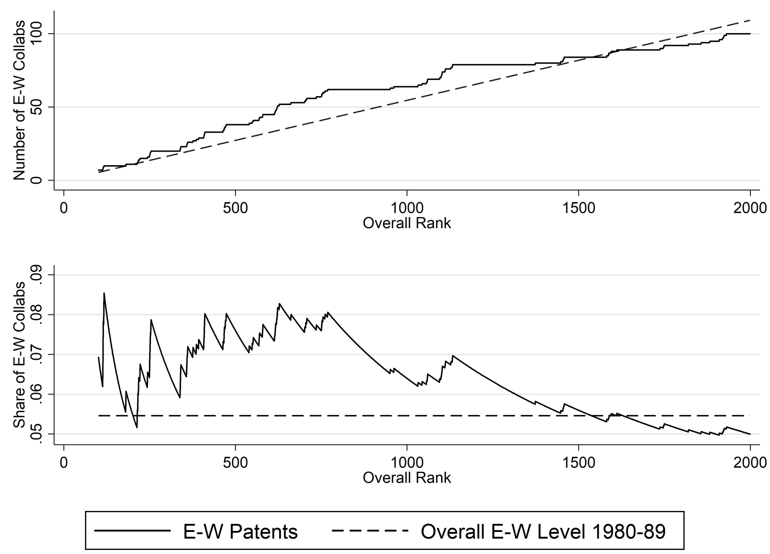
4 Identification Strategy

As the knowledge frontier keeps expanding, the complexity of innovation increases, often exceeding the capabilities of any single contributor. Consequently, collaborative interactions among inventors become crucial to turning ideas into tangible economic progress (Akcigit et al., 2018; Bloom et al., 2020; B. F. Jones, 2009). Inventor teams are mainly driven by the synergies of complementarities in their expertise.[11] Collaboration between East and West German inventors (hereafter referred to as *East-West collaborations* for brevity) was not completely impossible before the reunification, but it was very costly, not only because all cross-border interactions between East and West Germany were severely restricted, but also because scientific collaborations had to be approved before they could even begin. Despite such high administrative costs, East German and West German scientists collaborated on academic research projects in the late 1970s and 80s that were highly influential (Chan et al., 2022). This is also the case for East-West collaborations in patenting. Such collaborations gave way to patents deemed highly valuable, revealed by the number of citations and renewals these patents register. The share of patents that embody East-West collaborations makes up 5.4% of all $246,000$ West German patents from 1980 to 1989. When we rank these patents from the most to the least cited, we find disproportionately more East-West collaborative patents in top rankings as documented in Figure 1.

When East-West collaborative patents are ranked together with all West German patents registered between 1980 and 1989, we find 38 and 64 East-West patents within the top 500 and 1000 most cited patents, respectively. As shown in the lower panel of Figure 1, 38 and 64 East-West patents correspond to about 8% and 6.5% of the top 500 and top 1000 most cited patents, respectively, a larger share than the overall 5.4%, which is the overall share of East-West collaborations.

We determine complementarities between East and West German know-how that go into patenting by analyzing the IPC coverage of patents that embody an East-West collaboration. IPCs are indicators of the

Figure 1. Number and share of East-West collaborative patents in West German patents, ranked by citations

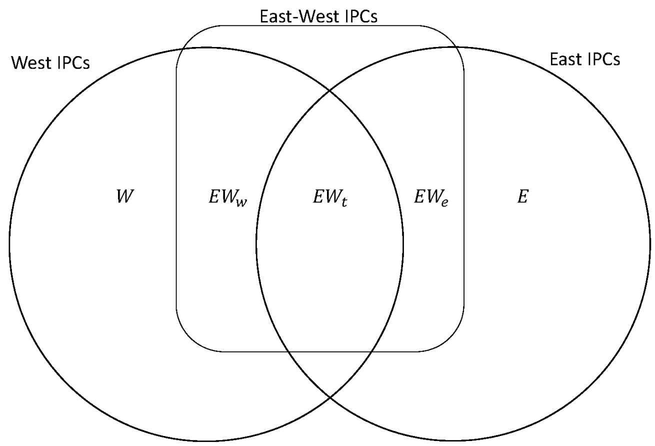


Notes: Citation rank of East-West German collaborative (E-W) patents in all West German patents that are registered from 1980 to 1989. Upper panel: The solid line shows the number of E-W patents (y-axis) in top X-ranked West German patents (x-axis). Lower panel: The solid line shows the share of E-W patents (y-axis) within the top X-ranked West German patents (x-axis). The dashed line in either panel shows what the number (upper panel) or share (lower panel) of E-W patents would be among the top X-ranked West German patents if the share of E-W patents in this range corresponded exactly to their average share in all West German patents.

content of specific scientific knowledge that feeds into patenting activities. Following Ferrucci (2020), we identify the top 75% of patent IPCs at the main group level, which is the fourth hierarchical level in the IPC taxonomy (WIPO, 2017), appearing on East German and West German patents between 1980 and 1989. We refer to the top 75% of IPCs among West German and East German patents as *West IPCs* and *East IPCs*, respectively. Similarly, we identify the top 75% of IPCs that appear in East-West collaborations between 1980 and 1989, which we refer to as *East-West IPCs*.[12] Figure 2 shows the overlaps between East, West, and East-West IPCs as follows: The area that remains inside West (East) IPCs but outside East-West IPCs in Figure 2, labeled W (E), contains West (East) IPCs that are not included among East-West IPCs. The area EW_w (EW_e) depicts those IPCs that are among West (East) IPCs as well as East-West IPCs but not among East (West) IPCs. The area EW_t depicts those IPCs that we find on patents of West Germany and East Germany as well as East-West collaborative patents.[13]

EW_e in Figure 2 depicts East German knowledge that goes into East-West collaborations and is not part of West German knowledge. The goal of the collaboration is to leverage the complementarities between the knowledge of the collaborators to achieve a result that neither party could have achieved alone (B. F. Jones, 2009), hence there must be some West German know-how that successfully complements EW_e to yield impactful patents. We find 1096 unique IPCs in East-West collaborative patents between 1980 and 1989, 655

Figure 2. Graphical representation of the identification strategy by IPC frequency category



Notes: The figure illustrates the identification strategy, categorizing each IPC according to its observed frequency across East, West, and East-West patents. The area EW_w and EW_e shows the top West and East German IPCs of East-West collaborations. EW_t represents the area of common IPCs on East and West German patents and East-West German collaborative patents.

of which appear among West and East IPCs. This area is labeled EW_t in Figure 2. 245 of the 1096 IPCs are East IPCs but not West IPCs, hence the area EW_e in Figure 2. When we investigate East-West collaborations from 1980 to 1989 to find out which strictly West IPCs (from the area EW_w in Figure 2) are listed in conjunction with these 245 IPCs, we identify 179 West IPCs. For traceability reasons, we refer to them as complementary West IPCs or, in short, complementary IPCs. These 179 complementary IPCs embody the West German know-how that was either non-existent or very scarce in East Germany and were used in combination with East German know-how to create East-West patents during times when such collaborations were very costly.

Based on the patents registered in each West German county between 1980 and 1989, we calculate the share of complementary IPCs in these counties. We calculate this share by dividing the number of patents with at least one complementary IPC by the total number of patents. This measure is the identification for our analysis. We argue that West German counties with a larger share of patents containing complementary IPCs are more likely to benefit from the reunification since these counties possess essential West German knowledge to complement the East German expertise that would enter West Germany after the reunification. Neither the fall of the Berlin Wall nor the reunification of Germany was an event that could be anticipated until the autumn of 1989 at the latest. Hence, no county could have strategically invested in these specific IPCs in the 1980s, anticipating intense future collaborations between East German and West German inventors (Hoisl et al., 2016). Figure A2 in the Appendix depicts the share of complementary IPCs in West German counties' patent registrations from 1980 to 1989.

5 Patenting Activities and Economic Outcomes: Results

In this section, we investigate how regional innovation affects individuals' labor market outcomes. Subsection 5.1 shows the results of the knowledge complementarities on the innovation outcome. In subsection 5.1, we derive the aggregate effects of regional innovation on income, duration of unemployment and probability of self-employment (avoiding reverse causality problems) in subsection 5.2. In subsection 5.3, we disentangle the aggregate effects and account for heterogeneity in how innovation affects the income of individuals employed in the local innovation sectors.

5.1 Collaborations and inventor Mobility as Drivers of Economic Outcomes

East-West German Collaborations. Our identification strategy is based on the expectation that a county's pre-reunification patenting experience paves the way for even more valuable innovative activity in the post-reunification period. If a West German county's pre-reunification patents imply the existence of local knowledge that best complements the East German knowledge that went into the East-West collaborations of the 1980s, we expect this county to be able to attract highly valuable East German knowledge after the reunification. One way to verify whether our identification correctly captures the likelihood of such counties being able to tap into East German know-how in the post-reunification era is to check for difference-in-differences in East-West collaborations. We estimate a linear probability model for patents registered in West German counties that embody East-West collaborations from 1985 to 1996, where the treatment is the share of complementary IPCs in a county, and the treatment period starts with the reunification in 1990.[14]

We regress a binary variable indicating whether a patent p registered in county c in year t embodies an East-West collaboration (EW_{pct}) on the interaction of the treatment and the treatment period to estimate

$$EW_{pct} = \beta(Complementary\ IPCs_c \times post90_t) + \mathbf{X}'_{pct}\zeta + \eta_c + \eta_t + \varepsilon_{pct} \quad (1)$$

where $\mathbf{X}'_{pct}$ is a vector of patent-level control variables containing the patent-specific number of inventors and the indicators for technological classes (derived from the IPCs classes). Variable descriptions and summary statistics are shown in Tables A1 and A14, respectively, in Appendix A. The main coefficient of interest is β. Positive and significant difference-in-differences would indicate a greater likelihood of West German counties with complementary IPCs registering East-West collaborations after the reunification.

As an alternative specification, we calculate the share of patents involving an East-West collaboration in all patents registered in county c in year t ($ShareEW_{ct}$) and use this county level share as the dependent variable to estimate

$$ShareEW_{ct} = \gamma(Complementary\ IPCs_c \times post90_t) + \mathbf{X'}_{ct}\zeta + \eta_c + \eta_t + \varepsilon_{ct} \quad (2)$$

Columns (1) and (2) in Table 1 report coefficient estimates for β in Equation (1). Columns (3) and (4) in Table 1 report coefficient estimates for γ in Equation (2). A one standard deviation increase in the share of complementary IPCs in a county's pre-reunification patent portfolio is associated with a 26% of one standard deviation increase in the share of East-West collaborations registered in that county. Counties harboring strong complementarities to pre-reunification East German knowledge have positive and significant difference-in-differences in the share of East-West collaboration patents after the reunification. Year and treatment interactions remain statistically insignificant during the pre-reunification period but are significant for years after the reunification, as shown in Figure A5. Hence, the parallel trend assumption is satisfied in both cases.

Table 1. East-West Collaborations and Complementary IPCs in West German Counties

	Patent Level: Likelihood of East-West Collaboration		County Level: Share of East-West Collaboration	
	(1)	(2)	(3)	(4)
Complementary IPCs × post90	0.408^a	0.395^a	0.541^a	0.543^a
	[0.0588]	[0.0309]	[0.0936]	[0.0622]
Controls	Yes	Yes	Yes	Yes
County FE	No	Yes	No	Yes
Std Errors	county	county × year	county	robust
Obs.	635270	635270	3238	3238
R2	0.0934	0.0995	0.0470	0.354

Notes. All specifications contain year-fixed effects. Patent-level controls are the number of inventors and technology class indicators. County-level controls are the total number of patents registered.
$^c\ p < 0.10$, $^b p < 0.05$, $^a\ p < 0.01$

East-West knowledge complementarities may capture not only West German-specific knowledge but general complementarities between East German technologies and the West, including Western Europe and the United States. To specify East-West German complementarities and to show that our identification is not driven by overall complementarities between East German and Western technologies, we run the estimates based on Equations (1) and (2) using an alternative definition of West German complementary knowledge, namely removing those IPCs that are also common to USPTO-registered patents. The difference-in-differences effects are shown in Table A6 in the Appendix.[15] The results are qualitatively similar to those in Table 1, thus confirming that our results are driven by local complementarities within West German counties rather than by complementarities between East German and Western or US technologies.

Mobility of East German Inventors. High-skilled migration enables knowledge diffusion (Kerr et al., 2017; Lissoni, 2018; Williams, 2006). We further argue, based on previous literature (Miguelez and Morrison, 2023; Putterman and Weil, 2010), that this also has an impact on economic outcomes. Next, we identify East German inventors who move to West German counties by tracking their locations of patent applications. We create an indicator variable to show whether a patent has an East German inventor on its list of inventors who migrated to a West German county before patent p is published. Similar to Equation (1), we use a linear probability model to estimate difference-in-differences for the likelihood that a patent p in county c in year t lists an East German inventor who migrated to that county. A positive and significant difference-in-differences effect would indicate that after the reunification, counties with a high share of complementary IPCs produce relatively more East-West collaborative patents, but these patents contain contributions of East German inventors who migrated to West Germany after the reunification. In an alternative specification, based on Equation (2), we investigate county-level effects by estimating difference-in-differences in the ratio of East German inventors residing in a West German county in a year to all inventors in that county and year.

Table 2. Emigrating East German Inventors and Counties with Complementary IPCs

	Patent Level: Likelihood of Emigrant on Patent		County Level: Share of Emigrants in County	
	(1)	(2)	(3)	(4)
Complementary IPCs $\times$ post90	0.434^{a}	0.424^{a}	0.287^{a}	0.286^{a}
	[0.0501]	[0.0281]	[0.0523]	[0.0384]
Controls	Yes	Yes	Yes	Yes
County FE	No	Yes	No	Yes
Std Errors	county	county $\times$ year	county	robust
Obs.	635270	635270	3218	3218
R2	0.0592	0.0656	0.173	0.447

Notes. All specifications contain year-fixed effects. Patent-level controls are the number of inventors and technology class indicators. County-level controls are the total number of patents registered.
^{c} $p < 0.10$, $^{b}p < 0.05$, ^{a} $p < 0.01$

As shown in Table 2, we obtain positive and significant difference-in-differences effects in both specifications. There is a relatively higher likelihood that a patent registered in a county with a high share of complementary IPCs in its pre-reunification patent portfolio will have an East German inventor who migrated to that West German county. A one standard deviation increase in the share of complementary IPCs in a county's pre-reunification patent portfolio is associated with a 44% of one standard deviation increase in the share of East German migrant inventors in that county. Interactions of year fixed effects with treatment are plotted in Figure A7.

West German counties with a high share of complementary IPCs are well positioned to facilitate cooperation and attract East German inventors, which stimulates regional patenting activities. This finding paves the way for the causal analysis of the effects of regional innovation on regional economic outcomes in the next subsection. In particular, the inflow of East German know-how is an important component of local patenting activity so that we can credibly claim that innovation precedes other movements in economic fundamentals. Moreover, the underlying exclusion restriction of our identification strategy, that counties with very low complementarities to East German know-how do not benefit as much from reunification is confirmed.

5.2 County Economic Performance

In this subsection, we analyze the causal effects of regional innovation on regional economic outcomes, which we capture via individual income, duration of unemployment, and the likelihood of being self-employed. We run difference-in-differences analysis with the treatment based on the identification presented above. The stable unit treatment value assumption (SUTVA) is crucial for the difference-in-differences, implying that spillover effects should be ruled out. The existence of complementarities in counties to East German patenting may not be spatially independent because counties with a high share of complementary IPCs tend to be regionally clustered, as shown in Figure A2. As a result of this spatial structure of the treatment, we would underestimate the difference-in-differences effects rather than overestimating them if post-reunification patenting activity in a county has spillover effects on neighboring counties. It is important to recall that spillover effects are geographically highly localized (Audretsch and Feldman, 1996; Feldman and Audretsch, 1999; Jaffe et al., 1993) and the effect of patenting in one region on another region declines sharply as the distance between them increases (Bottazzi and Peri, 2003). We estimate the following baseline Equation:

$$EconOutcome_{ict} = \delta(Complementary\ IPCs_{ic} \times post90_t) + \mathbf{X'}_{ict}\zeta + \eta_i + \eta_t + \varepsilon_{ict} \tag{3}$$

where $Complementary\ IPCs_{ic}$ is defined as in previous subsections, where the treatment variable is the proportion of the fractional count of patents registered under the complementary IPCs. The leading coefficient of interest, δ, is the effect of the difference in innovative potential differences across West German counties on economic outcomes. $EconOutcome_{ict}$ denotes the economic outcome for individual i residing in county c in year t where we capture the outcome using three proxies, namely the logarithm of individual i's labor income, the logarithm of the number of months by the total number of months in a year where i was unemployed, and the probability of self-employment. $\mathbf{X'}_{ict}$ is a vector of control variables. The terms η_i and η_t are individual and time fixed effects, and ε_{ict} is the error term clustered at the county level. The regressions are weighted using cross-sectional survey weights. This controls for bias in results due to oversampling of low-income households (Solon et al., 2015). We provide summary statistics for all outcomes in Table A14.

Throughout our estimations, we control for the interaction between the distance to the inner German border and the year to account for the fact that the results are not driven by geographical characteristics but only by the knowledge contained in the complementary IPCs. Alternatively, the results may be influenced by the increasing relevance of the geographical proximity to the inner German border over the years. According to Lichter et al. (2021), agriculture, energy/mining, and textiles were the main interests of the East German authorities when the district boundaries were drawn in 1952. We argue that controlling for the occupation classes of employed individuals in these industries in West Germany is also important because it allows us to control for possible changes in West German income driven by the migration of East Germans from these main East German industries to West Germany. We also control for the share of East German inventors who migrated to West Germany relative to the total number of inventors in the county. Overall, our estimation results reveal that individuals based in counties with stronger complementarities to pre-reunification East German knowledge have experienced increased labor income and self-employment.

Table 3. Economic Performance and Counties with Complementary IPCs

	Income		Months Unemployment		Self-employment	
	(1)	(2)	(3)	(4)	(5)	(6)
Complementary IPCs × post90	0.344^c	0.395^c	0.028	0.017	0.104	0.107
	[0.2040]	[0.2115]	[0.0587]	[0.0592]	[0.0878]	[0.0900]
Controls	No	Yes	No	Yes	No	Yes
Obs.	26626	26626	39742	39742	27120	27120
R2	0.8205	0.8205	0.4331	0.4338	0.6936	0.6938

Notes. The models are conducted at the individual-year level. We use the logarithmic transformation for the current gross income and months of unemployment variables. All regressions consider individual and year-fixed effects and are weighted using individual cross-sectional weights. Standard errors are clustered at the county level and provided in parentheses.
$^c\ p < 0.10$, $^b p < 0.05$, $^a\ p < 0.01$

A one standard deviation increase in a county's pre-reunification complementarity to East German knowledge, corresponding to a 3.7 percentage point increase in complementarity, increases its residents' income by 1.3% to 1.5% on average. Although the effect on income seems economically small, we emphasize that this is the average effect through all occupations covered in our SOEP sample. In the next subsection, we isolate the effects on incomes from occupations associated with local innovation sectors.

5.3 Occupation - Technology Proximity and Labor Market Outcomes

Our findings of the previous subsection are county-level averages. However, the literature points out the heterogeneous effects of innovation across income levels and occupations. This is primarily due to innovations heterogeneous effects on productivity and wages (Mann and Püttmann, 2023; Rocchetta et al., 2022).

To address such heterogeneity and establish possible causal links empirically, we explore how the overlap of a county's occupation portfolio with its patent portfolio affects labor market outcomes. Moreover, in our research design, differences in county-level complementary knowledge could have been determined by the prevailing occupation structure in a region and respective differences in income (Breau et al., 2014).

We capture the convergence of occupation and patent portfolios using a proximity indicator between the occupations of SOEP individuals and technology embodied in patent registrations in the respective county and year. Let P_{ct} indicate a patent registered in county c in year t. We follow the literature and apply a fractional count of patents for IPCs, which allocates in our setting proportional parts of equivalent NACE divisions to the patent (Brusoni et al., 2005; De Rassenfosse et al., 2013; Ferrucci, 2020). This way, the same patent is not assigned to different NACE divisions more than once. In addition, by considering all patent classes available within the patent application, we can argue that our proximity measure remains unbiased and accurate (Benner and Waldfogel, 2008). Each patent is attributed the respective fractional count value for a NACE division τ resulting in $\frac{1}{n_\tau}$. The total fractional count of NACE divisions within a given county and year is captured by vector $[Z_{ct}(1), \ldots, Z_{ct}(\tau), \ldots, Z_{ct}(\mathrm{T})]$ where $Z_{ct}(\tau) = \sum_{P_{ct}} \frac{1}{n_\tau}$. An individual in the SOEP data has only one occupation at a time, translating into a single NACE division so that the corresponding vector for the individual is $M_{i(c)t}(\tau)$. We compute an "exposure" measure of the NACE division of individual i's occupation to the patents' NACE vector of the residence county c in year t. We denote this as the $Occupation - Technology\ Proximity_{i(c)t}$.

We report in Table A9 the OLS coefficient estimates obtained from regressing the labor market outcomes of individual i residing in county c in year t on the proximity of this individual's occupation and the county's patent portfolio. We obtain a positive and statistically significant association between income and proximity measures. A one standard deviation increase in the proximity measure is associated with a 2% increase in income according to the OLS coefficient reported in column (1) of Table A9. This suggests better labor market outcomes in regions where patenting sectors are the main employers. However, OLS coefficients not only capture economically small average effects, but they also indicate a mere correlation. To overcome this issue, we use a two-stage least squares estimation to run a difference-in-differences estimation similar to previous subsections for occupation and patent portfolios' proximity in the first stage. We construct our instrument as follows

$$Occupation - \widehat{Technology}\ Proximity_{i(c)t} = \widehat{\delta}(Complementary\ IPCs_c \times post90_t) \tag{4}$$

where $Occupation - \widehat{Technology}\ Proximity_{i(c)t}$ are the predicted values of the exposure of an individual's occupation to the patent portfolio of the respective county. Our findings in previous sections show that West German counties with pre-existing complementarities to East German know-how benefit most from the German reunification. Therefore, we expect these counties to foster innovation and maintain and create employment in sectors where complementary know-how is widely used. The exclusion restriction holds since this is a plausible mechanism by which a county's pre-reunification complementarities to East German know-how could affect its post-reunification labor market outcomes. In the second stage, we regress individuals' income, duration of unemployment, and likelihood of self-employment on the instrumented occupation-technology proximity. Our two-stage least squares (2SLS) analysis recovers the disparity in outcomes between treated and untreated units by eliminating group and period effects (Gardner, 2022).

Table 4. Technological Proximity - 2SLS estimates

	Income		Months Unemployment		Self-employment	
	(1)	(2)	(3)	(4)	(5)	(6)
Occupation-Technology Proximity	3.015^{b}	3.203^{b}	0.121	0.036	1.169^{c}	1.333^{b}
	[1.3175]	[1.4342]	[0.4865]	[0.5168]	[0.6026]	[0.6606]
Cragg-Donald F	79.500	69.510	57.993	50.786	78.103	68.706
Kleibergen-Paap LM	21.912	19.196	15.304	13.613	21.668	19.069
	[0.0000]	[0.0000]	[0.0001]	[0.0002]	[0.0000]	[0.0000]
First-stage						
Complementary IPCs $\times$ post90	0.127^{a}	0.121^{a}	0.094^{a}	0.091^{a}	0.123^{a}	0.119^{a}
	[0.0291]	[0.0298]	[0.0258]	[0.0262]	[0.0285]	[0.0293]
First stage F-stat	18.903	16.551	13.438	11.953	18.723	16.468
Controls	No	Yes	No	Yes	No	Yes
Obs.	21062	21062	25146	25146	21385	21385

Notes. The estimations are conducted at the individual-year level, considering individual and year-fixed effects. We use the logarithmic transformation for the current gross income and months of unemployment variables. All regressions are weighted using individual cross-sectional weights. Robust standard errors are in parentheses.
^{c} $p < 0.10$, ^{b} $p < 0.05$, ^{a} $p < 0.01$

Table 4 reports the results from the estimation Equation (3) with the 2SLS estimator. Moreover, the coefficients offer insights into how pre-existing knowledge complementarities have affected the labor market success of individuals working in occupations with high occupation-technology proximity to East German knowledge. As we only consider the SOEP individuals within a county, this effect cannot be caused by a change in occupations due to the migration of SOEP individuals from West Germany to East Germany or vice versa.

Our 2SLS results reveal a robust and statistically significant impact of the instrumented occupation-technology proximity measure on economic performance. As revealed by Cragg-Donald's F statistic and Kleibergen-Paap's LM statistic, diagnostic tests imply that the proximity is being reliably instrumented. A one standard deviation increase in the county's IPC portfolio's complementarity is associated with an increase of 4% to 6% of one standard deviation in the respective proximity measure. The instrumented proximity measures reveal a statistically and economically significant effect on income. Precisely, a one standard deviation increase in the instrumented proximity increases income by 27% to 29%. The difference in OLS and 2SLS estimation results can be attributed to local average treatment effects captured by the 2SLS, which are larger than the overall average effects captured by the OLS. While the OLS coefficients capture the average effect of being in an occupation closely aligned with a county's patent portfolio, the 2SLS coefficients capture the effect of such alignment in a county with strong complementarities to East German know-how. Thus, the income of SOEP individuals in counties with high complementary knowledge are more likely to benefit from East German know-how after the reunification as their occupations align closer with sectors that experience a relative increase in patenting activity after the reunification. The effect of the proximity measure on the duration of unemployment is still a tightly estimated zero. However, the instrumented proximity measure is less noisy and statistically significantly positive. In the next section, we validate the robustness of our results.

6 Robustness Analysis

This section presents selected robustness checks for the results provided in the previous section. First, we assess the robustness of our results provided in section 5.1 by considering the pre-reunification availability of East German IPCs in West German counties that do not appear among East-West collaborations. This way, we argue that the increase in East-West collaborations in West German counties after the reunification is driven mainly by complementarities between the East and West German IPCs and not merely by the existence of East German knowledge in West German counties before the reunification.

Our identification strategy relies on complementarities of West and East German knowledge embodied in pre-reunification East-West collaborative patents. A plausible consideration is that the East German knowledge that went into the East-West collaborations before the reunification may have already existed in West German counties. When we examine the pre-reunification West German patents for top East and non-top West IPCs, which we previously identified as crucial East German contributions to pre-reunification collaborative patents, we find that such IPCs exist in the West but are very scarce. On average, they account for about 3% of the IPC portfolio in a West German county. We find 37 West German counties where the share of these IPCs is greater than 5%, as opposed to 224 West German counties where the share of complementary West German IPCs is greater than 5%. As anticipated, the top East German IPCs contributing to East-West collaborative patents represent scarce knowledge in West Germany. Furthermore, the correlation between the share of East German IPCs and complementary West German IPCs in any given West German county is very low (0.0065) and statistically insignificant. We rerun specifications (2) and (4) of Table 1 and specifications (2) and (4) of Table 2 with the share of East German IPCs in the respective county's IPC portfolio during the 1980s as a control variable. The estimated coefficients are reported in the first two and last two columns, respectively, of Table A7 in the Appendix. We find difference-in-differences effects consistent with those in Tables 1 and 2, thereby resolving our concerns.

Second, we report the estimated coefficients on the leading coefficient of interest δ in Equation (3) in Table A10.[16] Positive and statistically significant point estimates for δ suggest that the differences in labor income are changing in favor of counties with greater complementarities to the East German technologies as these experienced a strong inflow of complementary East German know-how after the reunification to boost local innovation, as shown in previous subsections. We find no significant effect on the duration of unemployment as this gamma-in-differences effect is a tightly estimated zero. However, we find a positive yet noisy difference-in-differences effect on the likelihood of being self-employed. While the point estimate of β exceeds its standard error, indicating a potentially positive effect that may be subject to considerable noise, it does not reach conventional significance levels by standard criteria.

Finally, we consider an alternative level of analysis, the labor market regions. For analysis of the labor market outcome of SOEP individuals, we use the information on the labor market regions from IAB Arbeitsmarktregionen (Bundesagentur fr Arbeit: Arbeitsmarktregionen (IAB), accessed July 10, 2024). While counties, or NUTS3 regions, represent administrative areas, labor market regions function as economically compact spatial units. Using this analysis level, we minimize potential spatial bias and correctly account for commuter flows (Kropp and Schwengler, 2011; Mewes and Broekel, 2020). Therefore, when aggregating the county-level data from SOEP to the labor market regions, the probability of commuters across labor market regions decreases (e.g., around 10% of employees commute across regions as compared to the almost 40% of commuting across counties (Kropp and Schwengler, 2011)). Figure A8 plots the geographical distribution of the complementary IPCs based on the count of patent applications, number of forward citations, and normalized number of patent renewals. In comparison to the geographical distribution at the county level, we observe a variation in the distribution of values across labor market regions, with no clear clustering pattern. This allows us to examine whether our results in the previous section are driven by the clustering effects of innovation in selected administrative regions.

We first perform the analysis from section 5.2 at the labor market regions. Table A15 in the Appendix reports the results. The labor market regions with a higher share of complementary IPCs experience no significant effect on the duration of unemployment and a very noisy positive effect (but do not fail to qualify as statistically significant) on the probability of self-employment. Surprisingly, the results for the impact on the income of individuals are not statistically significant for the share of complementary IPCs based on the number of patents. However, we find positive and statistically significant results when considering the share of complementary IPCs based on the number of citations and the normalized number of patent renewals. This indicates that for the impact of the income in the labor market regions, the quality rather than quantity of the patent is of higher relevance. Furthermore, when we employ more conservative fixed effects, such as labor market region fixed effects, we even find a significantly positive impact of the normalized number of renewals on the probability of self-employment. However, due to the restrictive definition of our measure of self-employment, which does not include further information on duration, these results must be viewed with caution.

Finally, we perform the analysis of the occupation-technology proximity at the labor market region level. Table A16 summarizes the OLS and 2SLS results, where occupation-technology proximity is instrumented by the pre-existing knowledge complementarity on the labor market outcomes of individuals. We affirm the robust positive effects on the income of individuals in a labor market region working in an occupation that is consistent with the patent portfolio of the respective region based on its pre-existing knowledge complementarity. The results on the duration of unemployment and the probability of self-employment are consistent with those presented in the previous section.

7 Conclusion

This study provides a causal analysis of the effects of patenting activities on individuals' income, duration of unemployment, and the likelihood of being self-employed. Using the natural experiment of German reunification to establish causality in a difference-in-differences setting, our analysis shows that individuals labor income in West German counties with strong complementarities to the East increased significantly. A one standard deviation increase in a county's pre-reunification complementarity to East German knowledge increases its residents' income by 1.3% to 1.5% on average. We estimate a positive but noisy difference-in-differences effect on individuals self-employment probability, and we find zero effect on the duration of unemployment.

Since these estimates are county-level average effects, we compute an annual proximity measure between each occupation and the countys patent portfolio to account for heterogeneous effects. We regress individuals income, unemployment, and the likelihood of self-employment on the instrumented proximity by pre-existing knowledge complementarities. Our results show that a one standard deviation increase in the instrumented proximity increases income by 27% to 29%. The results for the probability of self-employment are positive and significant. However, the duration of unemployment remains unaffected. We perform a series of robustness checks to validate our results.

We contribute to the existing literature by connecting different strands of literature on the impact of innovation, especially knowledge complementarities, on regional economic development. Our findings emphasize the significance of leveraging regional know-how to attract thriving research cooperation. With growing regional economic disparities threatening to undermine long-term economic outcomes, these issues become increasingly important.

Based on the findings of this analysis, we identify the following important avenues for future research. First, while we have carefully localized the patent data, advances in text analysis techniques for patent data may provide opportunities to correct missing information on patents that were previously unattributed and, therefore, excluded from our data analysis. Second, patenting activity indicates technological progress, and in our setting, it allows us to identify the effects of technological progress on regional labor market outcomes. However, the tacit knowledge that has migrated from the eastern part of Germany to the western part, which could not be captured here due to data limitations, may play a decisive role in shaping the long-term development of a region.

Notes

[1]We feed 40 years of data into our name-matching algorithm to identify individual inventors. We use data from 1980 to 1989 to calculate the share of certain technologies in regional patents. Our main analysis is based on data from 1984 to 1996.

[2]We use this term with a certain level of caution. We identify the East and West German-based inventors according to the location they indicate on the patent application.

[3]Evidence on the East German espionage in West German confirms the significance of knowledge transfer to East Germany (Glitz and Meyersson, 2020)

[4]Kogut and Zander (2000) show in the example of Carl Zeiss, a company specializing in optical instruments, which had to split to East and West Germany after WWII, that the technological profiles of both companies remained remarkably similar.

[5]Previous literature has identified various reasons for migration to West Germany, including social ties (Hoisl et al., 2016), economic perspectives (Fuchs-Schündeln and Schündeln, 2009), and political views (Rainer and Siedler, 2009).

[6]The analysis of patent renewals at the European Patent Office (EPO) and the US Patent and Trademark Office (USPTO) is based on data on patent life extension payments. For the US, there are three points within the patent term at which renewals can be made (in the 4th, 8th, and 12th year). In Europe, an annual fee applies from the second year after filing. For East and West Germany, the identification of patent renewals is based on the negative definition, i.e., the lack of payment. In West Germany, these payments were due in the third year, while in East Germany, they were due in the fourth year. Nowadays, a patent term of 20 years from filling is common in all countries. There are also exceptions: in individual cases and for specific industrial sectors, some patents run longer and can be extended more frequently. For all analyses involving renewals, the dataset is restricted to the four authorities mentioned above.

[7]PATSTAT provides information on the technology classification based on the International Patent Classification (IPC) and the Cooperative Patent Classification (CPC). Because of the reduced information on the CPC classes for East German patents, we focus on the IPC classification throughout our analysis.

[8]Through this thorough localization of patents, we lose ca. 20% from our "raw" dataset, where most of the loss is attributed to the misspelling of the postal code or missing additional information regarding, e.g. the city, which could allow us to identify a city in East or West Germany in case of a misspelled postal code.

[9]While SOEP data starts in 1984, information on the county-level data is available beginning in 1985.

[10]Note that the local innovation sector is defined by a patent's IPC code, not by the namesake variable $sector$ in the PATSTAT table $tls206\_person$, which indicates whether an entity is, for example, a university, company, or hospital.

[11]We use knowledge, know-how, and expertise interchangeably to denote the intangible assets utilized in patenting activities.

[12]For simplicity, we use the primary category without granular subclassifications (e.g., G01H1 for G01H1/00) in our analysis to capture the general technology aspects.

[13]Furthermore, Figure 2 correctly shows that we find some IPCs in East-West patents that neither belong to West German nor East German top 75% IPCs.

[14]We leave the year of the reunification 1990 and the following year out of this and all subsequent analyses because East Germany continued to exist until October 1990, while there were no restrictions on cross-border interactions and there were many administrative uncertainties surrounding the German reunification. Cross-border travel became possible with the opening of the border between East and West Germany on November 9, 1989. We do not exclude 1989 from our analysis because it would be unrealistic to expect that East-West patenting cooperation, which had not been planned before the opening of the border, would flourish during this period.

[15]Year and treatment interactions are shown in Figure A6 in the Appendix, which cautions us about county-level analysis using non-US complementarities. Therefore, our initial identification is a more appropriate treatment for the analysis at hand.

[16]We adopt the methodology as outlined in section 4 and derive the share of patent forward citations and normalized number of patent renewals based on complementary IPCs. Considering these measures of patent quality, we perform the regression specified in Equation 3 and present the results in Table A8. Additionally, Table A11 and Table A12 show the results for the lagged analysis.

Acknowledgements

We thank Alexander Kemnitz, Christian Lessmann, Korneliusz Pylak, Sunica Vuji and the participants of Regional Studies Association's Student and Early Career Conference (Cambridge, UK), 15th International SOEP Conference (Berlin, Germany), DRUID 2024 Conference (Nice, France), Regional Studies Associations Annual Conference 2024 (Florence, Italy), 2nd Workshop on the Economics of Science and Innovation (Bordeaux, France), 7th Geography of Innovation Conference (Manchester, UK), University of Portsmouth Research Seminar (Portsmouth, UK), Brown Bag Seminar at CEPIE, TUD Dresden University of Technology (Dresden, Germany), UB Summer School 2023 (Barcelona, Spain), 1st REGIS Summer School 2023 (Pisa, Italy) for their valuable comments and intensive discussion that significantly improved the paper. We would also like to express our gratitude to the SOEPremote team at DIW Berlin for their valuable support. Ali Sina Önder gratefully acknowledges financial support from the Institute for Humane Studies (grant no. IHS016734). An earlier version of this paper circulated under the title "*Innovation Patterns and Regional Development*".

References

Acemoglu, D. (2021). Harms of AI. *NBER Working Paper*, *No. 29247*.

Acemoglu, D., and Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy*, *128*(6), 2188–2244.

Acs, Z. J., Anselin, L., and Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy*, *31*(7), 1069–1085.

Aghion, P., Akcigit, U., Bergeaud, A., and Blundell, R. (2019). Innovation and top income inequality. *Review of Economic Studies*, *86*, 1–45.

Aghion, P., and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, *60*(2), 323–351.

Aghion, P., Howitt, P., and Prantl, S. (2015). Patent rights, product market reforms, and innovation. *Journal of Economic Growth*, *20*, 223–262.

Agrawal, A., Goldfarb, A., and Teodoridis, F. (2016). Understanding the changing structure of scientific inquiry. *American Economic Journal: Applied Economics*, *8*(1), 100–128.

Akçomak, I. S., and ter Weel, B. (2009). Social capital, innovation and growth: Evidence from Europe. *European Economic Review*, *53*, 544–567.

Akcigit, U., Caicedo, S., Miguelez, E., Stantcheva, S., and Sterzi, V. (2018). Dancing with the stars: Innovation through interactions. *NBER Working Paper*, *No. w24466*.

Akcigit, U., Grigsby, J., and Nicholas, T. (2017). Immigration and the rise of American ingenuity. *American Economic Review*, *107*(5), 327–331.

Asheim, B. T. (2012). The changing role of learning regions in the globalizing knowledge economy: A theoretical re-examination. *Regional Studies*, *46*(8), 993–1004.

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

Autor, D. H., and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, *103*(5), 1553–1597.

Avdeenko, A. (2018). Long-term evidence of retrospective voting: A natural experiment from the German Democratic Republic. *European Economic Review*, *103*, 83–107.

Azoulay, P., Fons-Rosen, C., and Zivin, J. S. G. (2019). Does science advance one funeral at a time? *American Economic Review*, *109*(8), 2889–2920.

Beblo, M., and Görges, L. (2018). On the nature of nurture. The malleability of gender differences in work preferences. *Journal of Economic Behavior & Organization*, *151*, 19–41.

Benner, M., and Waldfogel, J. (2008). Close to you? Bias and precision in patent-based measures of technological proximity. *Research Policy*, *37*(9), 1556–1567.

Bessen, J. (2015). *Learning by doing: The real connection between innovation, wages, and wealth*. Yale University Press.

Bilbao-Osorio, B., and Rodríguez-Pose, A. (2004). From r&d to innovation and economic growth in the eu. *Growth and Change*, *35*(4), 434–455.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, *110*(4), 1104–1144.

Bloom, N., Van Reenen, J., and Williams, H. (2019). A toolkit of policies to promote innovation. *Journal of economic perspectives*, *33*(3), 163–184.

Borjas, G. J., and Doran, K. B. (2012). The collapse of the soviet union and the productivity of American mathematicians. *The Quarterly Journal of Economics*, *127*(3), 1143–1203.

Borjas, G. J., and Doran, K. B. (2015). Which peers matter? The relative impacts of collaborators, colleagues, and competitors. *The Review of Economics and Statistics*, *97*(5), 1104–1117.

Boschma, R. A., and Fritsch, M. (2009). Creative class and regional growth: Empirical evidence from seven European countries. *Economic Geography*, *85*(4), 391–423.

Bottazzi, L., and Peri, G. (2003). Innovation and spillovers in regions: Evidence from European patent data. *European Economic Review*, *47*(4), 687–710.

Breau, S., Kogler, D. F., and Bolton, K. C. (2014). On the relationship between innovation and wage inequality: New evidence from canadian cities. *Economic Geography*, *90*(4), 351–373.

Brosig-Koch, J., Helbach, C., Ockenfels, A., and Weimann, J. (2011). Still different after all these years: Solidarity behavior in East and West Germany. *Journal of Public Economics*, *95*(11-12), 1373–1376.

Brusoni, S., Criscuolo, P., and Geuna, A. (2005). The knowledge bases of the worlds largest pharmaceutical groups: what do patent citations to non-patent literature reveal? *Economics of Innovation and New Technology*, *14*(5), 395–415.

Brynjolfsson, E., and McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Brynjolfsson, E., Mitchell, T., and Rock, D. (2018). What can machines learn and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, *108*, 43–47.

Bundesamt, S. (2022). *Gemeindeverzeichnis-Informationssystem GV-ISys.* Retrieved from https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/_inhalt.html# (accessed: 13-08-2024)

Callaert, J., Du Plessis, M., Grouwels, J., Lecocq, C., Magerman, T., Peeters, B., ... Vereyen, C. (2011). Patent statistics at eurostat: Methods for regionalisation, sector allocation and name harmonisation. *Eurostat Methodologies and Working Papers*.

Capello, R., and Lenzi, C. (2013). Territorial patterns of innovation and economic growth in European regions. *Growth and Change*, *44*(2), 195–227.

Capello, R., and Lenzi, C. (2019). Regional innovation evolution and economic performance. *Regional Studies*, *53*(9), 1240–1251.

Chan, H., Lariviere, V., Moy, N., Önder, A. S., Schilling, D., and Torgler, B. (2022). Science after communism: Why does westernization correlate with productivity? *SSRN*, *No. 3961712*.

Crescenzi, R., and Rodríguez-Pose, A. (2011). *Innovation and regional growth in the European Union*. Springer Science & Business Media.

Del Monte, A., Moccia, S., and Pennacchio, L. (2020). Regional entrepreneurship and innovation: Historical roots and the impact on the growth of regions. *Small Business Economics*, *58*, 451–473.

De Rassenfosse, G., Dernis, H., and Boedt, G. (2014). An introduction to the patstat database with example queries. *Australian Economic Review*, *47*(3), 395–408.

De Rassenfosse, G., Dernis, H., Guellec, D., Picci, L., and De La Potterie, B. V. P. (2013). The worldwide count of priority patents: A new indicator of inventive activity. *Research Policy*, *42*(3), 720–737.

EUROSTAT. (2001). Verordnung (EG) Nr. 29/2002 der Kommission vom 19. Dezember 2001 zur nderung der Verordnung (EWG) Nr. 3037/90 des Rates betreffend die statistische Systematik der Wirtschaftszweige in der Europischen Gemeinschaft. *Amtsblatt der Europischen Gemeinschaften*, *L6*(3), 33.

Feldman, M. P., and Audretsch, D. B. (1999). Innovation in cities: Science-based diversity, specialization and localized competition. *European Economic Review*, *43*(2), 409–429.

Ferrucci, E. (2020). Migration, innovation and technological diversion: German patenting after the collapse of the Soviet Union. *Research Policy*, *49*(9), 104057.

Frey, C. B., and Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerization? *Technological Forecasting and Social Change*, *114*, 254–280.

Fritsch, M. (2001). Co-operation in regional innovation systems. *Regional Studies*, *35*(4), 297–307.

Fritsch, M., Greve, M., and Wyrwich, M. (2023). Shades of a socialist legacy? Innovation activity in East and West Germany 1877-2014. *University of Groningen, FEB Research Institute*.

Fritsch, M., and Wyrwich, M. (2021). Does successful innovation require large urban areas? Germany as a counterexample. *Economic Geography*, *97*(3), 284–308.

Fuchs-Schündeln, N., and Schündeln, M. (2009). Who stays, who goes, who returns? East–West migration within Germany since reunification. *Economics of Transition*, *17*(4), 703–738.

Gardner, J. (2022). Two-stage differences in differences. *arXiv*, *2207.05943*.

Glitz, A., and Meyersson, E. (2020). Industrial espionage and productivity. *American Economic Review*, *110*(4), 1055-1103.

Goebel, J. (2014). Job submission instructions for the SOEPremote system at DIW Berlin. *SOEP Survey Papers*, *195*.

Goebel, J., Grabka, M. M., Liebig, S., Kroh, M., Richter, D., Schröder, C., and Schupp, J. (2019). The German socio-economic panel (SOEP). *Jahrbücher für Nationalökonomie und Statistik*, *239*(2), 345–360.

Grabka, M. M. (2024). SOEP-Core v38 codebook for the PEQUIV file 1984-2020: CNEF variables with extended income information for the SOEP. *SOEP Survey Papers*.

Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature*, *28*(4), 1661–1707.

Hall, B., Jaffe, A. B., and Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools. *NBER Working Paper*, *No. 8498*.

Hall, B., and van Reenen, J. (2000). How effective are fiscal incentives for R&D? A review of the evidence. *Research Policy*, *29*(4-5), 449–469.

Harhoff, D., Narin, F., Scherer, F. M., and Vopel, K. (1999). Citation frequency and the value of patented inventions. *Review of Economics and Statistics*, *81*(3), 511–515.

Hegde, D., Herkenhoff, K., and Zhu, C. (2023). Patent publication and innovation. *Journal of Political Economy*, *131*(7), 1845-1903.

Heineck, G., and Süssmuth, B. (2013). A different look at lenins legacy: Social capital and risk taking in the two germanies. *Journal of Comparative Economics*, *41*(3), 789–803.

Hémous, D., and Olsen, M. (2022). The rise of the machines: Automation, horizontal innovation, and income inequality. *American Economic Journal: Macroeconomics*, *14*(1), 179–223.

Hnermund, P., and Hipp, A. (2024). *Inventor mobility after the fall of the berlin wall.* Retrieved from https://arxiv.org/abs/2409.01861

Hoisl, K., Harhoff, D., Dorner, M., Hinz, T., and Bender, S. (2016). Social ties or patent quality signals-evidence from east german inventor migration. In *Academy of management proceedings* (Vol. 2016, p. 13594).

Hunt, J., and Gauthier-Loiselle, M. (2010). How much does immigration boost innovation? *American Economic Journal: Macroeconomics*, *2*(2), 31–56.

Iammarino, S., Rodríguez-Pose, A., and Storper, M. (2019). Regional inequality in europe: evidence, theory and policy implications. *Journal of Economic Geography*, *19*(2), 273–298.

Jaffe, A. B., and Lerner, J. (2011). *Innovation and its discontents: How our broken patent system is endangering innovation and progress, and what to do about it*. Princeton University Press.

Jaffe, A. B., and Trajtenberg, M. (2002). *Patents, citations, and innovations: A window on the knowledge economy*. MIT press.

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, *108*(3), 577–598.

Jones, B. F. (2009). The burden of knowledge and the death of the renaissance man: Is innovation getting harder? *The Review of Economic Studies*, *76*(1), 283–317.

Jones, C. I. (2023). Recipes and economic growth: A combinatorial march down an exponential tail. *Journal of Political Economy*, *131*(8), 1994-2031. Retrieved from https://doi.org/10.1086/723631 doi: 10.1086/723631

Jones, C. I., and Kim, J. (2018). A schumpeterian model of top income inequality. *Journal of Political Economy*, *126*(5), 1785–1826.

Kerr, S. P., Kerr, W., Özden, Ç., and Parsons, C. (2017). High-skilled migration and agglomeration. *Annual Review of Economics*, *9*(1), 201–234.

Kogan, L., Papanikolaou, D., Schmidt, L. D. W., and Seegmiller, B. (2021). Technology, vintage-specific human capital, and labor displacement: Evidence from linking patents with occupations. *NBER Working Paper*, *No. 29552*.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics*, *132*(2), 665–712.

Kogut, B., and Zander, U. (2000). Did socialism fail to innovate? A natural experiment of the two zeiss companies. *American Sociological Review*, *65*(2), 169–190.

Kropp, P., and Schwengler, B. (2011). Abgrenzung von Arbeitsmarktregionen – ein Methodenvorschlag. *Raumforschung und Raumordnung*, *69*(1), 45–62.

Laudenbach, C., Malmendier, U., and Niessen-Ruenzi, A. (2020). The long-lasting effects of living under communism on attitudes towards financial markets. *NBER Working Paper*, *No. 26818*.

Lichter, A., Löffler, M., and Siegloch, S. (2021). The long-term costs of government surveillance: Insights from stasi spying in East Germany. *Journal of the European Economic Association*, *19*(2), 741–789.

Lissoni, F. (2018). International migration and innovation diffusion: An eclectic survey. *Regional Studies*, *52*(5), 702–714.

Mann, K., and Püttmann, L. (2023). Benign effects of automation: New evidence from patent texts. *Review of Economics and Statistics*, *105*(3), 562–579.

Maraut, S., Dernis, H., Webb, C., Spiezia, V., and Guellec, D. (2008). The OECD REGPAT database: A presentation. *OECD Science, Technology and Industry Working Papers*, *2008*(02), 1–36.

Mewes, L., and Broekel, T. (2020). Technological complexity and economic growth of regions. *Research Policy*, *51*(8), 104156.

Miguelez, E., and Morrison, A. (2023). Migrant inventors as agents of technological change. *Journal of Technology Transfer*, *48*(2), 669–692.

Möhlmann, A. (2014). Persistence or convergence? The EastWest tax-morale gap in Germany. *Public Finance Analysis*, 3–30.

Mokyr, J., Vickers, C., and Ziebarth, N. L. (2015). The history of technological anxiety and the future of economic growth: Is this time different? *Journal of Economic Perspectives*, *29*(3), 31–50.

Moser, P., Ohmstedt, J., and Rhode, P. W. (2018). Patent citationsAn analysis of quality differences and citing practices in hybrid corn. *Management Science*, *64*(4), 1926-1940.

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

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press.

Önder, A. S. (2022). Ostdeutsche Wissenschaft nach der Wende: Institutioneller Umbau und Produktivität. *ifo Dresden berichtet*, *29*(04), 23–28.

Prantl, S., and Spitz-Oener, A. (2020). The impact of immigration on competing natives' wages: Evidence from German reunification. *Review of Economics and Statistics*, *102*(1), 79–97.

Putterman, L., and Weil, D. N. (2010). Post-1500 population flows and the long-run determinants of economic growth and inequality. *The Quarterly Journal of Economics*, *125*(4), 1627–1682.

Rainer, H., and Siedler, T. (2009). Does democracy foster trust? *Journal of Comparative Economics*, *37*(2), 251–269.

Rocchetta, S., Ortega-Argilés, R., and Kogler, D. F. (2022). The non-linear effect of technological diversification on regional productivity: Implications for growth and smart specialisation strategies. *Regional Studies*, *56*(9), 1480–1495.

Rodríguez-Pose, A., and Crescenzi, R. (2008). Research and development, spillovers, innovation systems, and the genesis of regional growth in europe. *Regional studies*, *42*(1), 51–67.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*(5, Part 2), S71–S102.

Romer, P. M. (2009). Two strategies for economic development: using ideas and producing ideas. In *The strategic management of intellectual capital* (pp. 211–238). Routledge.

Sabel, B. A. (1993). Science reunification in Germany: A crash program. *Science*, *260*(5115), 1753–1758.

Sampat, B., and Williams, H. L. (2019). How do patents affect follow-on innovation? Evidence from the human genome. *American Economic Review*, *109*(1), 203-36.

Schmoch, U. (2008). Concept of a technology classification for country comparisons. *Final Report to the World Intellectual Property Organisation (WIPO)*.

SOEP. (2020). *SOEP-Core v35–PGEN: Person-related status and generated variables. SOEP survey papers 834: Series d variable descriptions and coding.* Berlin: DIW Berlin/SOEP.

Solon, G., Haider, S. J., and Wooldridge, J. M. (2015). What are we weighting for? *Journal of Human Resources*, *50*(2), 301–316.

Webb, M. (2019). The impact of artificial intelligence on the labor market. *SSRN 3482150*.

Weisskircher, M. (2020). The strength of far-right AfD in Eastern Germany: The east-west divide and the multiple becauses behind populism. *The Political Quarterly*, *91*(3), 614–622.

Weitzman, M. L. (1998). Recombinant growth. *The Quarterly Journal of Economics*, *113*(2), 331–360.

Williams, A. M. (2006). Lost in translation? International migration, learning and knowledge. *Progress in Human Geography*, *30*(5), 588–607.

WIPO. (2017). *Guide to the international patent classification* (Version 2017 ed.). Geneva, Switzerland: Author.

Wyrwich, M. (2019). Historical and current spatial differences in female labour force participation: Evidence from Germany. *Papers in Regional Science*, *98*(1), 211–240.

Appendix A Relevant Figures, Tables, and Results

A.1 Data

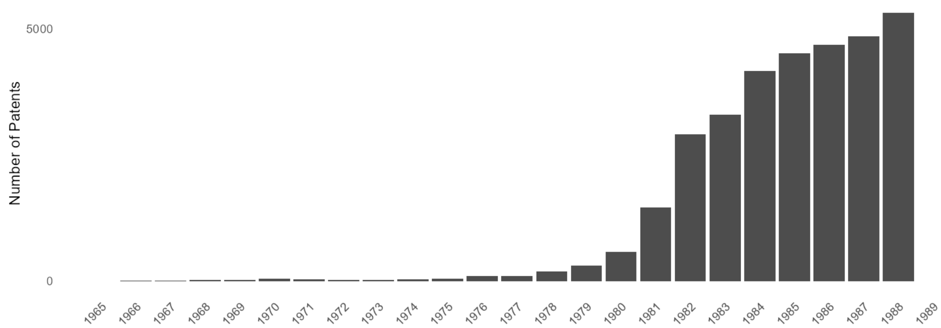
Table A1. Variable Description

Variable	Type	Level	Description	Source
Income	log	individual & year	Current gross labor income in Euro	GSOEP
Months Unemployment	log	individual & year	Number of moths of official unemployment/Total nmber of months in a calender year	GSOEP
Self-employment		individual & year	Probability of self-employment	GSOEP
Complementary IPCs		county	Share of West German patents on the top West German IPCs	PatStat
Complementary IPCs (Non-US)		county	Share of West German patents on the top West German but NonTop US IPCs	PatStat
Complementary IPCs Citations		county	Share of West German patent citations on the top West German IPCs	PatStat
Complementary IPCs Renewals		county	Share of West German patent normalized renewals on the top West German IPCs	PatStat
Patent EW		patent & year	Dummy East-West German collaborative patent	PatStat
County EW		county & year	Share of East-West German collaborative patent	PatStat
Patent Emigrant		patent & year	Dummy patent containing at least one East German emigrant as inventor	PatStat
County Emigrant		county & year	Share of patents containing at least one East German emigrant as inventor	PatStat
Number Inventors		patent & year	Aggregated number of inventors	PatStat
Technological Classes		patent & year	Patent-specific technological classes	PatStat
Number Patent Registrations	log	county & year	Aggregated number of patent registrations	PatStat
Share of Emigrant Inventors		county & year	Number of inventors that emigrated from East to West Germany divided by the total number of inventors	PatStat
Border Distance		county	Distance to the inner German border from 1989	Ahlfeldt
Employment Agriculture		county & year	Share of persons employed in agriculture in the total working age sample	GSOEP
Employment Energy/Mining		county & year	Share of persons employed in energy/mining industry in the total working age sample	GSOEP
Employment Textile		county & year	Share of persons employed in textile industry in the total working age sample	GSOEP
Occupation-Technology Proximity		individual & year	Exposure measure of Occupation to Patenting	Patstat & GSOEP

Notes. We account for inflation by calculating real income in 2000 prices. To ensure comparability of patent renewals across different patent authorities, we divide the number of renewals by the maximum number for a year and a patent authority. We apply the fractional count method as in Ferrucci (2020). Information on the distance to the inner German border is from https://github.com/Ahlfeldt/MRRH2018-toolkit/blob/5a1e05a7ba9f4f85218c4032a6b3de7b01a6a34c/data/input/CountyBorderDist.csv (accessed June 4, 2024)

A.2 PATSTAT Data: Descriptives

Figure A1. Number of distinct East-West German collaborative patents per year, 1965-1989



Notes: The figure plots the number of distinct East-West German collaborative patents per year. We define an East-West German collaborative patent as a collaboration between at least one inventor from a different country than the rest of the team.

Table A2. A Brief Sample of Correspondance of East and West IPCs from East-West Patents

East IPC	**Corresponding Complementary West IPC**
G01H1: Measurement of mechanical vibrations in solids by using conduction	F41J5: Weapons; target indication systems
F23D17: Burners for combustion simultaneously or alternately	F23Q7: Ignition using electrically-produced heat
H05H: Production of accelerated electrically-charged particles	H01J23: Details of transit-time tubes
C10G: Production of liquid hydrocarbon mixtures	B01J27: Catalysts comprising carbon compounds
C10G: Production of liquid hydrocarbon mixtures	C07C41: Preparation of ethers
C10G: Production of liquid hydrocarbon mixtures	F02B3: Engines characterized by air compression

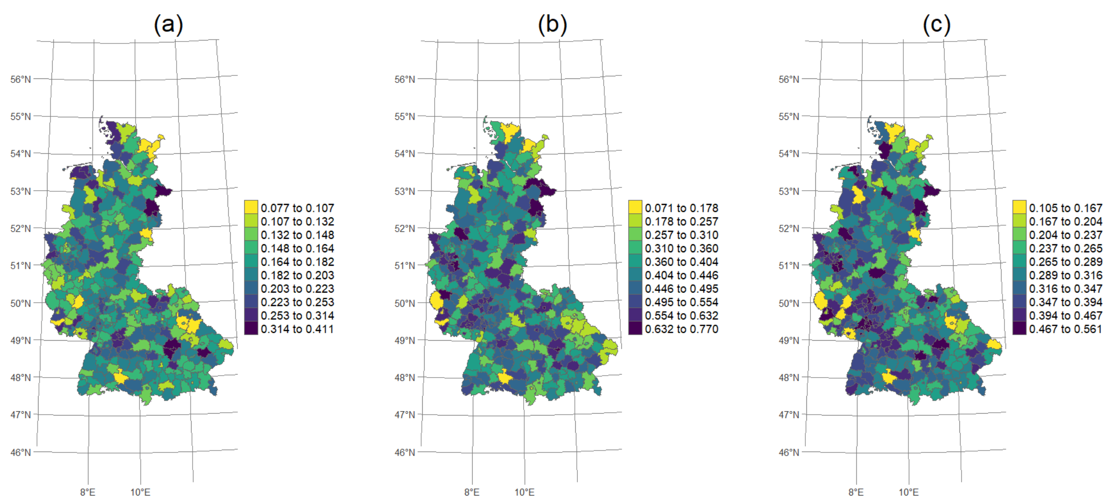
Notes. The table summarizes selected IPC examples considered in our sample.

Table A3. Top Technology Classes of Patents 1980-1989

Technology Class	Ranking (Share) within Patents		
	West Germany	East Germany	East-West German Collab
Chemistry: Organic fine chemistry	1 (7.2%)	15 (3.01%)	1 (7.3%)
Chemistry: Basic materials chemistry	2 (6.1%)	14 (3.04%)	2 (7.1%)
Electrical engineering: Electrical machinery	3 (5.4%)	4 (5.53%)	7 (4.86%)
Mechanical engineering: Mechanical elements	4 (5.1%)	7 (4.4%)	12 (3.3%)
Mechanical engineering: Other special machines	5 (5.01%)	5 (5.15%)	6 (4.9%)
Mechanical engineering: Transport	6 (5.0%)	13 (3.1%)	14 (2.9%)
Chemistry: Chemical engineering	7 (4.9%)	6 (4.9%)	5 (5.7%)
Chemistry: Macromolecular chemistry, polymers	8 (4.7%)	17 (2.6%)	10 (4.5%)
Instruments: Measurement	9 (4.5%)	1 (9.9%)	3 (6.8%)
Mechanical engineering: Handling	10 (4.04%)	3 (5.7%)	9 (4.6%)
Mechanical engineering: Machine tools	11 (4.02%)	2 (8.5%)	4 (5.8%)
Mechanical engineering: Textile or paper machines	12 (3.98%)	10 (3.4%)	11 (4.2%)
Mechanical engineering: Engines, pumps, turbines	13 (3.8%)	18 (2.5%)	20 (2.1%)
Civil engineering	14 (3.76%)	9 (3.7%)	15 (2.9%)
Chemistry: Materials, metallurgy	15 (3.0%)	8 (3.9%)	8 (4.6%)

Notes. The table summarizes the rankings and shares of the most popular technology classes among 1980s patents registered in either West or East Germany and among patents created by the two countries' collaborative efforts. Technology classes are determined based on Schmoch (2008). The share of a technology class is calculated based on the share of patents that report IPCs in line with that technology class. The ranking is based on shares.

Figure A2. Fractional count of patents, citations, renewals, 1980-1989



Notes: The figure plots the fractional count of (a) patents, (b) citations, and (c) renewals based on complementary IPCs for West German counties for the period from 1980 to 1989. Shapefile information is provided by the Federal Agency for Cartography and Geodesy (Gebietseinheiten 1:5 000 000 (GE5000), accessed July 20, 2024).

Figure A3. Example Patent Applications EPO

(19) **Europäisches Patentamt**
European Patent Office
Office européen des brevets

(11) Veröffentlichungsnummer: **0 279 312**
A2

(12) **EUROPÄISCHE PATENTANMELDUNG**

(21) Anmeldenummer: **88101799.0**

(51) Int. Cl.[4]: **C07C 76/06** , C07C 79/10

(22) Anmeldetag: **08.02.88**

---

(30) Priorität: **18.02.87 DE 3705091**

(43) Veröffentlichungstag der Anmeldung:
24.08.88 Patentblatt 88/34

(84) Benannte Vertragsstaaten:
BE DE ES FR IT

(71) Anmelder: **BAYER AG**
Konzernverwaltung RP Patentabteilung
D-5090 Leverkusen 1 Bayerwerk(DE)

(72) Erfinder: **Witt, Harro, Dr.**
Möhlenbarg 2
D-2224 Kuden(DE)
Erfinder: **Beckhaus, Heiko, Dr.**
Winterberg 17
D-5090 Leverkusen 3(DE)

Notes: Patent application at the European Patent Office. Based on the postal code indicated on the patent applications, we can identify the location of inventors, here in former West Germany.

Figure A4. Example Patent Applications in East Germany

(19) **DEUTSCHE DEMOKRATISCHE REPUBLIK**

PATENTSCHRIFT

Wirtschaftspatent

Erteilt gemaeß § 17 Absatz 1 Patentgesetz

ISSN 0433-6461 (11) **209 436**

Int.Cl.[3] 3(51) **C 07 C 79/12**
C 07 C 76/06

AMT FUER ERFINDUNGS- UND PATENTWESEN

In der vom Anmelder eingereichten Fassung veroeffentlicht

(21) WP C 07 C/ 2432 483 (22) 15.09.82 (44) 09.05.84

(71) siehe (72)
(72) SCHMIDT, FRANK,DR.-ING. DIPL.-ING.;BLAICH, LUTZ,DIPL.-ING.;HARTUNG, JUERGEN,DIPL.-ING.;
HAMANN, KORNELIA,DIPL.-CHEM.;DD;
WEISS, SIEGFRIED,PROF. DR. SC. TECHN.;WELSCH, GUENTER;
FISCHER, HORST,DR. RER. NAT. DIPL.-CHEM.;LINDAU, HORST;DD;

(54) **VERFAHREN ZUR ENERGIESPARENDEN TRENNUNG DES NITRO-O-DICHLORBENZENS VON DER RUECKSTANDSSAEURE DES NITRIERPROZESSES**

(57) Verfahren zur energiesparenden Trennung des Nitro-o-dichlorbenzens von der Rückstandssäure des Nitrierprozesses ohne Einsatz von Waschwasser. Es werden Koaleszenzhilfen zur schnellen und vollständigen Trennung vorgeschlagen. Die Erfindung dient der energiesparenden, die Umwelt nicht mit Nitro-o-dichlorbenzen-haltigen Abwasser belastenden Trennung des Nitro-o-dichlorbenzens von der Rückstandssäure. Das vorgeschlagene Verfahren ist mit kontinuierlich und diskontinuierlich betreibbaren Apparaturen realisierbar. Das Verfahren zeichnet sich durch verminderten Trennaufwand (geringer Energie-, kein Wasser- und Neutralisationsmittelbedarf) und verminderte Verweilzeit des zu trennenden Zwei-Phasengemisches in der Trennapparatur aus.

Notes: Patent application ("economic patent" (*Wirtschaftspatent*)) at the East German Patent Office. Based on the postal code indicated on the patent applications, we can identify the inventors' locations here in former East Germany.

A.3 SOEP Sample Restrictions and Composition based on County Information

Table A4. Sample Structure: Restriction and Inclusion Criteria based on County Information

	Observations (person-year)
Sample Construction	
All SOEP respondents from private households (1984-2000, Sample: A, E)	141.075
Attributing the County Information corresponding Complementary IPCs and Controls	106.817
Individual and Year Sample Restrictions	
Restrict to the period from 1985 to 1996	100.320
Retain only individuals appearing from 1985 to 1996	64.436
Exclude the years with uncertainties: 1990 and 1991	53.725
Main Variables Data Availability	
Current Gross Income	32.604
Months Unemployment	42.364
Self-Employment	29.123
Main Variables Sample Restriction	
Exclude the top and bottom 1% of the Income Variable	31.948
Restrict income to working age population	28.543
Final Sample	
Current Gross Income	26.834
Months Unemployment	39.871
Self-Employment	27.351

Notes. The table provides a detailed approach to the composition of the SOEP sample according to the labor market areas in Germany.

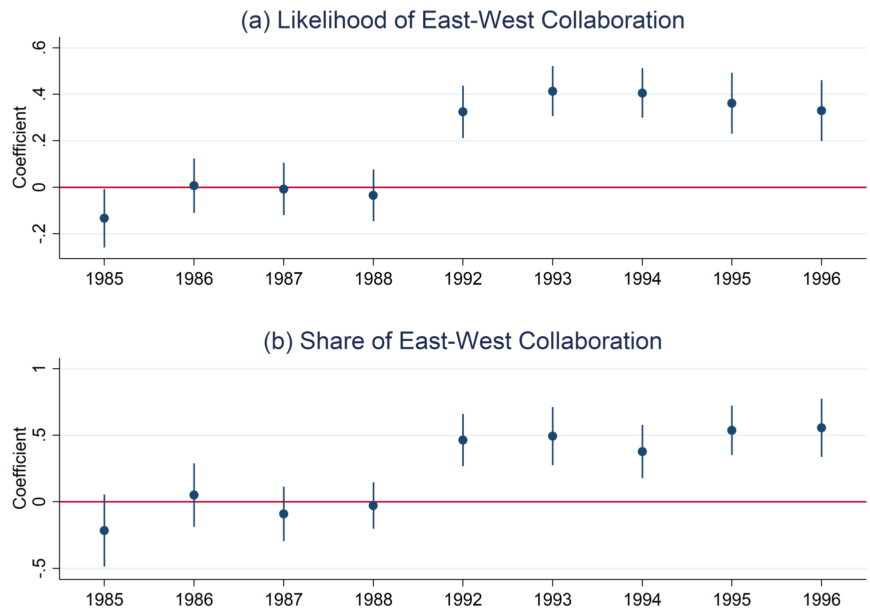
Table A5. Summary Statistics

	Mean	St. Dev.	p25	p50	p75	Min	Max	Obs.
Income	7.4327	0.6377	7.2126	7.5532	7.8143	5.2479	8.7178	26.590
Months Unemployment	0.0294	0.1240	0	0	0	0	0.6931	39.742
Self-employment	0.0520	0.2220	0	0	0	0	1	27.120
Complementary IPCs	0.0682	0.0365	0.0455	0.0596	0.0802	0.0073	0.2824	324
Complementary IPCs (Non-US)	0.0134	0.0109	0.0066	0.0104	0.0184	0	0.0735	324
Patent EW	0.0895	0.2855	0	0	0	0	1	40.3574
County EW	0.1094	0.0754	0.0625	0.1017	0.1408	0	0.6522	3.238
Patent Emigrant	0.0917	0.2885	0	0	0	0	1	40.3574
County Emigrant	0.0669	0.0589	0.0342	0.056	0.0859	0	0.75	3.218
Number Inventors	2.4148	1.7193	1	2	3	1	25	40.3574
Number Patent Registrations	4.555	1.252	3.7377	4.554	5.407	0.6931	7.8876	3.238
Overlap	0.0195	0.0802	0	0	0.003	0	0.9947	29304
Distance inner border	145.3390	76.7087	92.4760	155.9613	202.4676	0	324.0249	50.800
Share Inventors Emigrants	0.0484	0.0284	0.0345	0.0455	0.0584	0	0.5000	50.800
Employment Agriculture	0.0111	0.1048	0	0	0	0	1	50.800
Employment Mining	0.0112	0.1051	0	0	0	0	1	50.800
Employment Textile	0.0132	0.1143	0	0	0	0	1	50.800

Notes. Summary statistics for the main GSOEP variables. The variable Income and Months Unemployment are log-transformed

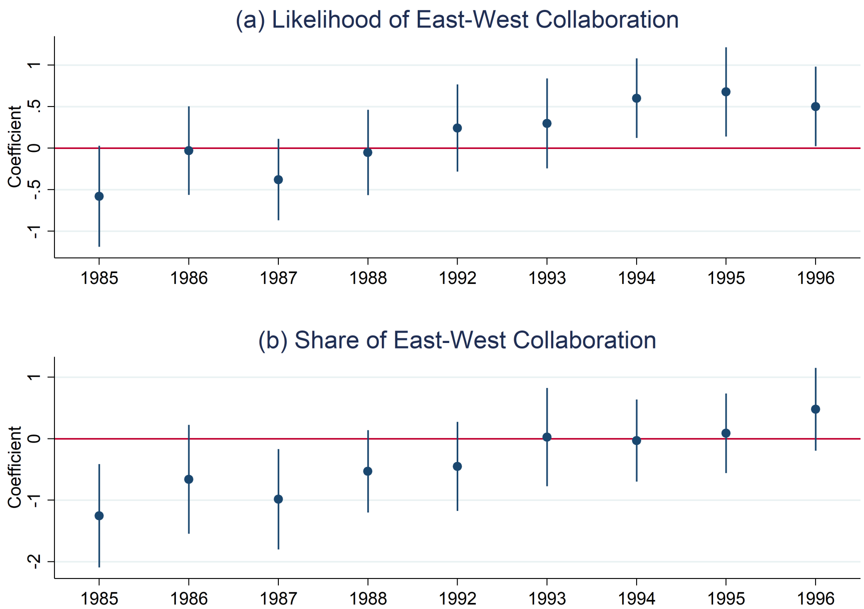
A.4 Additional Results

Figure A5. Year Interactions with Treatment (Complementarity to East German IPCs)



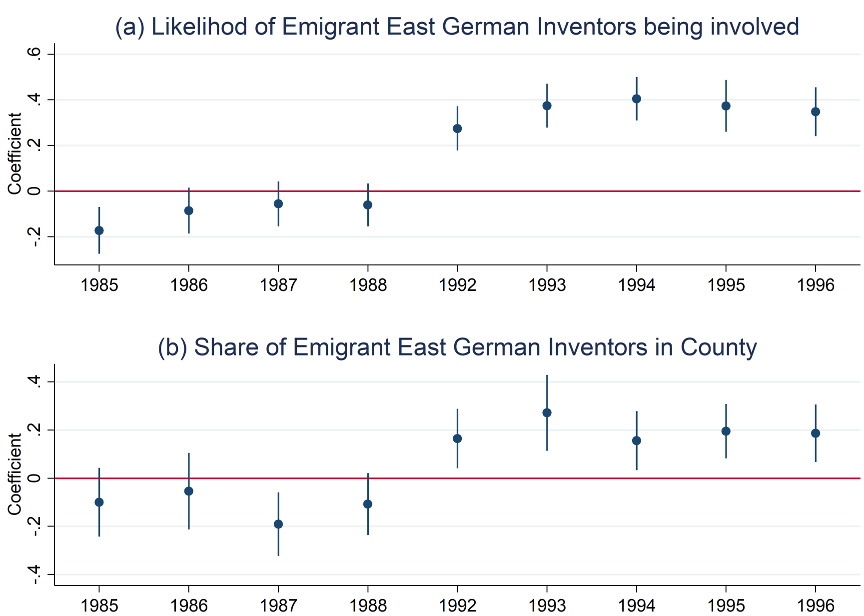
Notes: The figure plots the coefficients of the year and treatment interactions.

Figure A6. Year and treatment interactions for the identification excluding top US IPCs



Notes: The figure represents the coefficients of the year and treatment interactions.

Figure A7. Mobility of East German Inventors (Year Interactions with Treatment)



Notes: The figure represents the coefficients of the year and treatment interactions.

Table A6. Propensity and Share of East-West German Collaborations and top West German IPCs non-US IPCs

	Patent Level: Likelihood of East-West Collaboration		County Level: Share of East-West Collaboration	
	(1)	(2)	(3)	(4)
Complementary IPCs × post90	0.727^b	0.664^a	0.693^b	0.705^a
(w/o US Complementarity)	[0.340]	[0.145]	[0.299]	[0.204]
Controls	Yes	Yes	Yes	Yes
County FE	No	Yes	No	Yes
Std Errors	county	county × year	county	robust
Obs.	635270	635270	3238	3238
R2	0.0921	0.0989	0.0326	0.339

Notes. All specifications contain year-fixed effects. Patent-level controls are the number of inventors and technology class indicators. County-level controls are the total number of patents registered.
^{c} $p < 0.10$, $^b p < 0.05$, ^{a} $p < 0.01$

Table A7. East-West Collaborations and Complementary IPCs in West German Counties

	Collaboration		Mobility	
	Patent	County	Patent	County
Complementary IPCs × post90	0.408^a	0.541^a	0.435^a	0.287^a
	[0.0458]	[0.0689]	[0.0423]	[0.0441]
Share of East IPCs	0.214^a	0.247^a	0.188^a	-0.0424
	[0.0619]	[0.0926]	[0.0616]	[0.0608]
Std Errors	county × year	county × year	robust	robust
Obs.	635270	3238	635270	3218
R2	0.0934	0.0506	0.0592	0.173

Notes. All specifications contain year-fixed effects. Patent-level controls are the number of inventors and technology class indicators. County-level controls are the total number of patents registered. Whenever the share of East IPCs is included as a control, as these don't vary over the years, we drop the county fixed effects.
^{c} $p < 0.10$, $^b p < 0.05$, ^{a} $p < 0.01$

Table A8. Economic Performance and Counties with Complementary IPCs

	Income		Months Unemployment		Self-employment	
	(1)	(2)	(3)	(4)	(5)	(6)
Complementary Citations IPCs × post90	0.115^{c}	0.125^{b}	0.014	0.011	0.034	0.034
	[0.0597]	[0.0615]	[0.0166]	[0.0166]	[0.0245]	[0.0245]
Obs.	26626	26626	39742	39742	27120	27120
R2	0.8205	0.8206	0.4331	0.4338	0.6936	0.6938
Complementary Renewals IPCs × post90	0.157^{b}	0.0179^{b}	0.022	0.017	0.034	0.035
	[0.0707]	[0.0733]	[0.0198]	[0.0203]	[0.0313]	[0.0322]
Obs.	26626	26626	39742	39742	27120	27120
R2	0.8205	0.8207	0.4331	0.4338	0.6936	0.6938
Controls	No	Yes	No	Yes	No	Yes

Notes. The models are conducted at the individual-year level. We use the logarithmic transformation for the current gross income and months of unemployment variables. All regressions consider individual and year-fixed effects and are weighted using individual cross-sectional weights. Standard errors are clustered at the county level and provided in parentheses.
$^{c}\ p < 0.10$, $^{b}p < 0.05$, $^{a}\ p < 0.01$

Table A9. Technological Proximity - OLS estimates

	Income		Months Unemployment		Self-employment	
	(1)	(2)	(3)	(4)	(5)	(6)
Occupation-Technology Proximity	0.254^{a}	0.253^{a}	0.032	0.032	-0.066^{a}	-0.067^{a}
	[0.0692]	[0.0692]	[0.0255]	[0.0255]	[0.0159]	[0.0160]
Controls	No	Yes	No	Yes	No	Yes
Obs.	21062	21062	25146	25146	21385	21385
R2	0.8189	0.8189	0.3687	0.3687	0.7179	0.7179

Notes. The estimations are conducted at the individual-year level, considering individual and year-fixed effects. We use the logarithmic transformation for the current gross income and months of unemployment variables. All regressions are weighted using individual cross-sectional weights. Robust standard errors are in parentheses.
$^{c}\ p < 0.10$, $^{b}\ p < 0.05$, $^{a}\ p < 0.01$

Table A10. Economic Performance and Counties with Complementary IPCs

	One year		Two years		Three Years	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Income						
Complementary IPCs × post90	0.369^{c}	0.432^{b}	0.374^{c}	0.336	0.504^{c}	0.484^{c}
	[0.1942]	[0.2021]	[0.2262]	[0.2354]	[0.2676]	[0.2788]
Controls	No	Yes	No	Yes	No	Yes
Obs.	23060	23060	16882	16882	10898	10898
R2	0.8361	0.8363	0.8505	0.8506	0.8712	0.8713
Panel B: Months Unemployment						
Complementary IPCs × post90	0.024	0.020	-0.001	-0.034	-0.028	-0.068
	[0.0622]	[0.0630]	[0.0729]	[0.0744]	[0.0889]	[0.0894]
Controls	No	Yes	No	Yes	No	Yes
Obs.	34301	34301	25289	25289	16496	16496
R2	0.4557	0.4563	0.4947	0.4952	0.5637	0.5643
Panel C: Self-employment						
Complementary IPCs × post90	0.062	0.062	0.041	0.020	0.088	0.062
	[0.0819]	[0.0833]	[0.0875]	[0.0880]	[0.1062]	[0.1052]
Controls	-	Yes	-	Yes	-	Yes
Obs.	23506	23506	17215	17215	11141	11141
R2	0.7094	0.7096	0.7363	0.7365	0.7654	0.7657

Notes. The models are conducted at the individual-year level. We use the logarithmic transformation for the current gross income and months of unemployment variables. All regressions consider individual and year-fixed effects and are weighted using individual cross-sectional weights. Standard errors are clustered at the county level and provided in parentheses.
^{c} $p < 0.10$, $^{b}p < 0.05$, ^{a} $p < 0.01$

Table A11. Economic Performance and Counties with Complementary IPCs (based on Citations)

	One year		Two years		Three Years	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Income						
Complementary Citations IPCs × post90	0.120^c	0.135^b	0.130^c	0.118^b	0.159^c	0.155^c
	[0.0574]	[0.0587]	[0.0697]	[0.0710]	[0.0816]	[0828]
Controls	No	Yes	No	Yes	No	Yes
Obs.	23060	23060	16882	16882	10898	10898
R2	0.8361	0.8363	0.8505	0.8506	0.8713	0.8713
Panel B: Months Unemployment						
Complementary Citations IPCs × post90	0.013	0.013	0.012	0.004	0.012	0.003
	[0.0178]	[0.0176]	[0.0208]	[0.0214]	[0.0250]	[0.0253]
Controls	No	Yes	No	Yes	No	Yes
Obs.	34301	34301	25289	25289	16496	16496
R2	0.4558	0.4563	0.4948	0.4952	0.5637	0.5642
Panel C: SelfNoemployment						
Complementary Citations IPCs × post90	0.023	0.022	0.020	0.014	0.030	0.023
	[0.0226]	[0.0228]	[0.0248]	[0.0251]	[0.0304]	[0.0295]
Controls	No	Yes	No	Yes	-	Yes
Obs.	23506	23506	17215	17215	11141	11141
R2	0.7094	0.7096	0.7363	0.7363	0.7654	0.7657

Notes. The models are conducted at the individual-year level. We use the logarithmic transformation for the current gross income and months of unemployment variables. All regressions consider individual and year-fixed effects and are weighted using individual cross-sectional weights. Standard errors are clustered at the county level and provided in parentheses.
^{c} $p < 0.10$, $^b p < 0.05$, ^{a} $p < 0.01$

Table A12. Economic Performance and Counties with Complementary IPCs (based on Renewals)

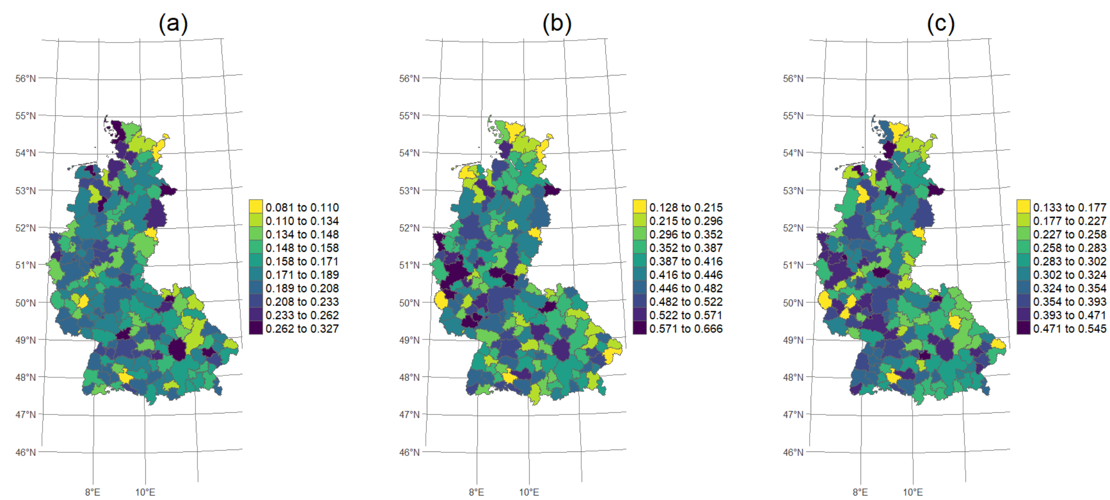
	One year		Two years		Three Years	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Income						
Complementary RenewalsIPCs × post90	0.167^{b}	0.195^{a}	0.182^{b}	0. 0.171^{c}	0.218^{b}	0.214^{b}
	[0.0686]	[0.0709]	[0.0840]	[0.0871]	[0.1007]	[0.1041]
Controls	No	Yes	No	Yes	No	Yes
Obs.	23060	23060	16882	16882	10898	10898
R2	0.8362	0.8363	0.8506	0.8506	0.8713	0.8713
Panel B: Months Unemployment						
Complementary Renewals IPCs × post90	0.021	0.020	0.017	0.006	0.015	0.000
	[0.0214]	[0.0218]	[0.0255]	[0.0267]	[0.0305]	[0.0314]
Controls	No	Yes	No	Yes	No	Yes
Obs.	34301	34301	25289	25289	16496	16496
R2	0.4558	0.4563	0.4948	0.4952	0.5637	0.5642
Panel C: Self-employment						
Complementary Renewals IPCs × post90	0.021	0.021	0.016	0.007	0.026	0.015
	[0.0294]	[0.0302]	[0.0307]	[0.0315]	[0.0356]	[0.0355]
Controls	No	Yes	No	Yes	No	Yes
Obs.	23506	23506	17215	17215	11141	11141
R2	0.7094	0.7096	0.7363	0.7363	0.7654	0.7657

Notes. The models are conducted at the individual-year level. We use the logarithmic transformation for the current gross income and months of unemployment variables. All regressions consider individual and year-fixed effects and are weighted using individual cross-sectional weights. Standard errors are clustered at the county level and provided in parentheses.
$^{c}\ p < 0.10$, $^{b}p < 0.05$, $^{a}\ p < 0.01$

A.5 Labor Market Regions

A.5.1 Descriptive Evidence

Figure A8. Fractional count of patents, citations, renewals, labor market areas, 1980-1989



Notes: The figure plots the fractional count of (a) patents, (b) citations, (c) renewals based on complementary IPCs for West German labor market areas ("Arbeitsmarktregionen"), for the period from 1980 to 1989. The patent data is attributed the labor market areas considering the information provided by Federal Institute for Research on Building, Urban Affairs and Spatial Development (Referenztabellen zu Raumgliederungen des BBSR). Shapefile information is provided by the Federal Agency for Cartography and Geodesy (Gebietseinheiten 1:5 000 000 (GE5000), accessed July 20, 2024).

A.5.2 Sample Restrictions and Composition based on LMR from IAB

Table A13. Sample Structure: Restriction and Inclusion Criteria based on LMR Information

	Observations (person-year)
Sample Construction	
All SOEP respondents from private households (1984-2000, Sample: A, E)	141.065
Attributing the Labor Market Regions	128.818
Attributing the Labor Market Regions corresponding Complementary IPCs	124.680
Individual and Year Sample Restrictions	
Restrict to the period from 1985 to 1996	100.325
Retain only individuals appearing from 1985 to 1996	64.461
Exclude the years with uncertainties: 1990 and 1991	53.746
Main Variables Data Availability	
Current Gross Income	32.624
Months Unemployment	42.383
Self-Employment	29.141
Main Variables Sample Restriction	
Exclude the top and bottom 1% of the Income Variable	31.968
Restrict income to working age population	28.561
Final Sample	
Current Gross Income	28.561
Months Unemployment	42.383
Self-Employment	29.141

Notes. The table provides a detailed approach to the composition of the SOEP sample according to the labor market areas in Germany.

Table A14. Summary Statistics LMR

	Mean	St. Dev.	p25	p50	p75	Min	Max	Obs.
Income	7.4414	0.6384	7.2134	7.5661	7.8280	5.2479	8.7178	28.974
Months Unemployment	0.0300	0.1245	0	0	0	0	0.6931	42.351
Self-employment	0.0506	0.2191	0	0	0	0	1	28.379

Notes. Summary statistics for the main SOEP variables LMR level. The variable Income and Months of Unemployment are log-transformed

A.5.3 LMR Results

Table A15. Economic Performance and LMR with Complementary IPCs

	Income		Unemployment		Self-employment	
	(1)	(2)	(3)	(4)	(5)	(6)
Complementary IPCs × post90	0.568	0.594	0.044	0.043	0.238	0.234
	[0.4330]	[0.4337]	[0.1305]	[0.1292]	[0.1436]	[0.1445]
Complementary IPCs Citations × post90	0.306^b	0.296^b	0.019	0.022	0.238	0.070
	[0.1442]	[0.1456]	[0.1305]	[0.0288]	[0.1436]	[0.0492]
Complementary IPCs Renewals × post90	0.335^b	0.326^b	0.054	0.054	0.238	0.105^b
	[0.1513]	[0.1515]	[0.1305]	[0.0441]	[0.1436]	[0.0444]
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
LMR Fixed Effects	No	Yes	No	Yes	No	Yes
Observations	28379	28379	42315	28974	42315	28974

Notes. The models are conducted at the individual-year level. We use the logarithmic transformation for the current gross income and months of unemployment variables. All regressions are weighted using individual cross-sectional weights. Standard errors are clustered at the county level and provided in parentheses.
^{c} $p < 0.10$, $^b p < 0.05$, ^{a} $p < 0.01$

Table A16. Technological Proximity - 2SLS estimates

	OLS			2SLS		
	Income	Months Unemployment	Self-Employment	Income	Months Unemployment	Self-Employment
	(1)	(2)	(3)	(4)	(5)	(6)
Occupation-Technology Proximity	0.041^a	-0.005	-0.016^a	0.559^b	0.088	0.156
	[0.0120]	[0.0037]	[0.0038]	[0.2413]	[0.0797]	[0.1169]
Cragg-Donald F				61.337	50.086	59.156
Kleibergen-Paap LM				54.934	50.109	53.826
				[0.0000]	[0.0001]	[0.0000]
Complementary IPCs × post90				1.124^a	0.876^a	1.091^a
				[0.1622]	[0.1312]	[0.1588]
First stage F-stat				47.994	44.585	47.138
Obs.	21080	24742	21527	21080	24742	21527

Notes. The models are conducted at the individual-year level, considering individual and year-fixed effects. We use the logarithmic transformation for the current gross income and months of unemployment variables. All regressions are weighted using individual cross-sectional weights. Robust standard errors are in parentheses. ^{c} $p < 0.10$, ^{b} $p < 0.05$, ^{a} $p < 0.01$

Appendix B German Socio-Economic Panel Data: Main Variables

Individual-level analysis is performed using the information from the SOEP (1984-2000). We use the remote-access version of the SOEP, SOEPremote, to create our sample of individuals and assign the information on the respondents' place of residence at the county level. Instructions on how to get access to SOEPremote as well as work with SOEPremote systems are provided in Goebel (2014) (https://hdl.handle.net/10419/96113). Access to the county-level information can be provided after a contract with SOEP/DIW (DIW Berlin: Regional Data). Therefore, SOEP data are not included in the replication files and can only be accessed via SOEPremote by users with a contract.

Current Gross Labor Income in Euro (pgen: pglabgro) We use the variable *pglabgro* variable, available from 1984. It represents "imputed gross labor income in the previous months for all SOEP respondents who are employed in a main job". The variable values exclude the one-time payments. In the case of this variable, missing values are imputed in a two-stage procedure.[17]

We use the information on the Consumer Price Index (CPI) provided by the Federal Statistical Office of Germany (*Statistische Bundesamt, Destatis*). Since our analysis is limited to West Germany, we apply the calculation for the *costs of living price index for all private households, former feral territory*. We use the year 2000 as the index year.

Officially Unemployed Previous Year Number Months (e.g., pkal:kal1d02) The duration of unemployment is calculated based on Lichter et al. (2021), where we divide the number of months registered as unemployed by the total number of months in the respective wave for each individual in each wave.

Occupational Position (pgen: pgstib) Since SOEP does not provide information on the months spent in self-employment per year. We derive a variable of the probability of an individual as well as the change from official employment to self-employment based on the information regarding the current occupational position. Provided that the individuals are currently of working age, we define self-employment when the current occupation is *"other self-employed", "Other self-employed, with 10 and more employees"*. The information on the current occupation is generated, according to SOEP, by combining the information on *"occupational group", "unemployed (yes/no)", "military/community service", "in education (yes/no)", "pensioner"*.

Industry sector of the current occupation (NACE Rev. 1.1, Division) (e.g., apgen: nace84) For employed individuals and each wave, SOEP provides information on the industry to which each individual's employment is attributed. This information is generated from the respondents' answers to the question about the industry in which they are currently employed. If no job change is reported, information from the previous year is imputed. The classification is according to the Statistical Classification of Economic Activities in the European Community (Nomenclature des statistiques des activits conomiques de la Communaut europenne - NACE Rev. 1.1), Level 3 (Division). As NACE Rev. 1.1 was the version used during the period of our analysis, we decided to use this version rather than a more recent one such as NACE Rev. 2.

Individuals Cross-sectional Weight without 1st wave of a subsample (pequiv: w11101) The variable provides information on the individual's population and sample weight. This information can be used to analyze income over time. The purpose of the weights is to account for unequal selection probabilities and sample attrition. These weights also include population weights to achieve population-representative results (Grabka, 2024). We use restriction criteria to include only individuals who are present for all years of our analysis. In this case, cross-sectional weights are applied.

Inverse staying probability Respondent Individuals (pequiv: w11103$$) To derive longitudinal weights, we multiply the individual's cross-sectional weight (w11101$$) with the inverse staying probability, which represents the probability that an individual is present in the respective wave (Grabka, 2024).

Appendix C PATSTAT

C.1 PATSTAT Data Processing Pipeline

Figure C1 illustrates the comprehensive data processing pipeline utilized for constructing the research dataset from the PATSTAT database. This process is critical for ensuring that the analysis is built on an accurately defined set of data. The process begins with the initial extraction of over 2 million patent records involving multiple tables. In particular, the dataset was defined using the following inclusion and exclusion criteria:

- All patent offices available in PATSTAT
- Only patents with a presence of an East or West German inventors: $person\_ctry\_code$ in table $tls206\_person$ equals "DD" or "DE"
- Only inventors with at least two patents between 1965 and 2004 (years according to variable $appln\_filing\_date$ in table $tls201\_appln$)
- Only priority filings: Only patents that have no prior application in table $tls204\_appln\_prior$
- Exclusion of utility models, PCT applications, provisional applications, design patents, plant patents, and artificial applications: $appln\_kind$ in table $tls206\_person$ equals "A"

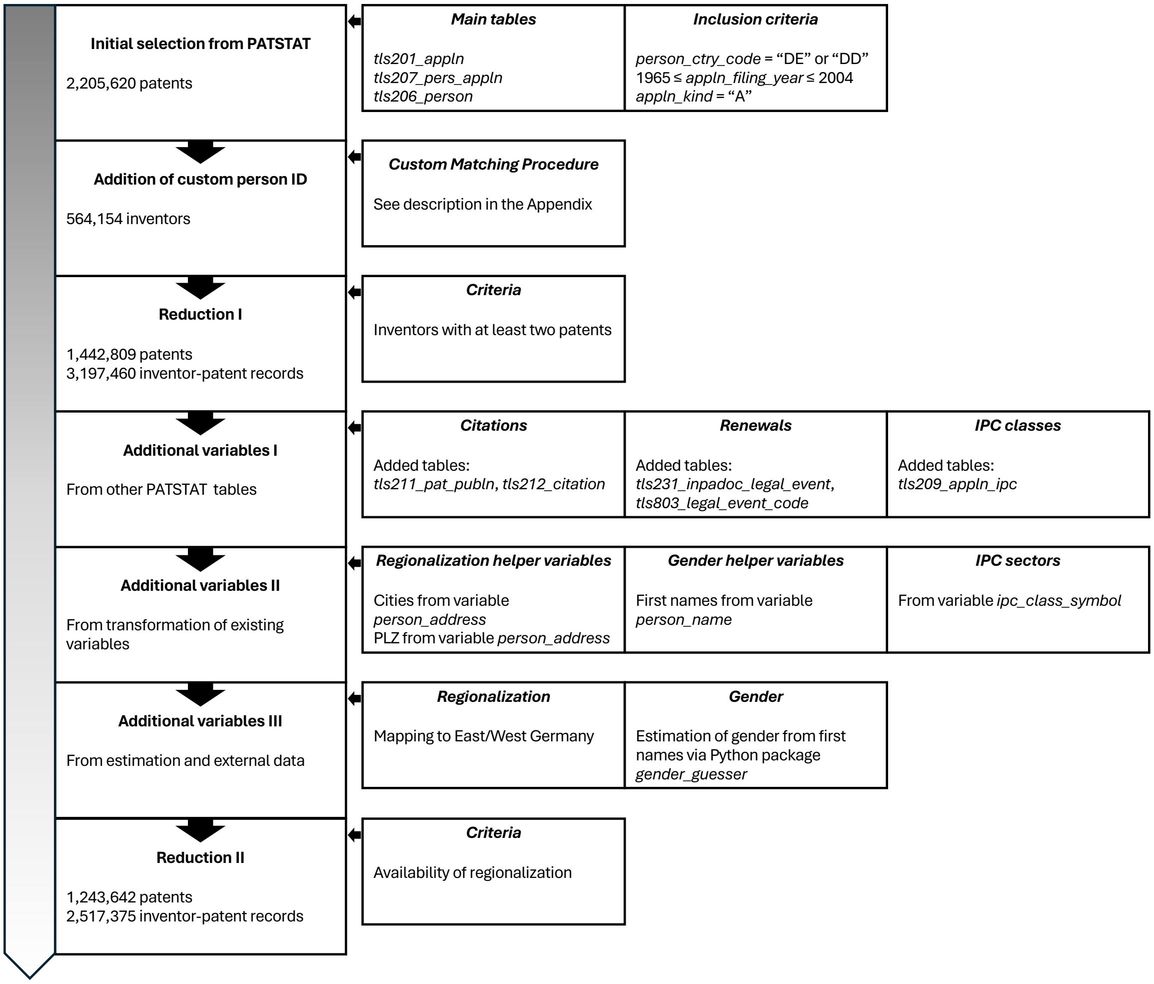
After the initial data extraction, a custom person ID is assigned to each inventor to ensure accurate tracking and matching across records. A detailed custom matching procedure, outlined in subsection C.3, is applied to ensure that only inventors with at least two patents are included, allowing for the analysis of repeated patent activity while excluding one-time filers. Additional variables are integrated from other PATSTAT tables, transformations of existing variables, and external data sources, including citation data, renewal information, and IPC classifications. Finally, the dataset is enhanced with regional information derived from inventor addresses and gender estimation based on first names.

C.2 PATSTAT Regionalization of East and West German Patents

In our analysis, we establish the regional connections of patents by utilizing the addresses provided by the inventors on their patent applications. Our study primarily concentrates on patent offices in East and West Germany. This specific geographic focus necessitated a deviation from using the OECD REGPAT Database (Maraut et al., 2008). The REGPAT Database provides regional linkage for patents filed with major offices such as the European Patent Office (EPO), the Patent Cooperation Treaty (PCT), and the United States Patent and Trademark Office (USPTO). However, its coverage does not extend to the detailed regional data required for our focus on German patent offices, leading us to adopt alternative methods for regional linkage in our research.

The original PATSTAT dataset includes NUTS3 level information, which pertains to the regionalization specifics of the Patstat data as detailed by Callaert et al. (2011). We have enhanced the dataset by retrieving postal code details from the address field of the patent applications, covering the years 1965 to 2004. Typically, this information is available in the variable $person\_address$ in table $tls226\_person\_orig$, and for some cases in the variable $person\_name$ in the same table in PATSTAT. Notably, during the period before the German reunification, West and East Germany maintained distinct postal systems, where codes sometimes overlapped. For instance, the postal code "8051" could refer to either "Dresden" in Saxony, East Germany, or "Allershausen, Oberbayern" in Bavaria, West Germany. In cases of such ambiguity, the dataset was refined by comparing with the location of an inventor's city after reunification. Manual checks were conducted where automatic resolution was inadequate.

Figure C1. Flowchart of the PATSTAT Data Processing Pipeline



Notes: The figure plots the PATSTAT Data Processing Pipeline.

After reunification in 1993, Germany adopted a uniform five-digit postal code system. In the absence of an official conversion table between the pre- and post-reunification codes, we sourced extensive postal code data from the website https://www.alte-postleitzahlen.de (accessed August 21, 2024) to align the old postal codes with the new ones, thus standardizing the dataset. Further, we converted the existing NUTS3 data into corresponding county codes using conversion tables provided by Bundesamt (2022). The refined dataset now provides information on both NUTS3 levels and county codes, linking them to both historical and current postal codes.

Additionally, we have complemented the NUTS3-based county code data with further county code information obtained by converting postal codes to county codes through the OpenStreetMap-based tool available at georef-germany-postleitzahl (accessed August 21, 2024). This process enables the attribution of county codes to each innovator based on the updated postal code data.

C.3 PATSTAT Construction of Inventor CVs

We derive a custom person ID for inventors to trace them by matching their names. Our disambiguation procedure utilizes a graph-theoretic and hierarchical method. Initially, we identified groups with identical surnames and constructed a graph depicting relationships among first names within these groups. First names were classified as identical, different, subsets, or partially compatible (e.g., "M. John" with "Michael J."). After establishing all relationships, we modeled these as nodes and edges in a graph, removed redundant connections, and analyzed non-forking subset paths from the graph's leaves upward. Due to the large number of names and the resulting high likelihood of arbitrary matches, we further required identical middle initials. In our framework, "Michael John" and "Michael J." with the same last name are matched if they are the sole variants, unlike when "Michael Jason" is also present, yielding three separate person IDs in this example. We did not consider addresses and other additional information as matching criteria for two reasons. First, the objective of the analysis is to establish CVs under the assumption of migration patterns. Second, the format and completeness of addresses in the PATSTAT database exhibit a high level of variation.

For each custom person ID, we determine their earliest known city of residence based on these IDs. Inventors whose city at the first time they appeared on a patent was located in an East/West German state were categorized as East/West German. To determine whether a city was located in the East or West, we mapped its zip code from the PATSTAT data to its state from the online data source postal-codes-json-xml-csv (accessed August 21, 2024). Due to its special role of a divided city, we excluded the city-state of Berlin from this analysis.

We estimate inventors' gender based on first names using the Python package $gender\_guesser$. Before feeding names to this algorithm, we removed the common German professional prefixes, suffixes, and titles Dr., Dipl. Ing. Chem., Biol. Phys., FH, Prof., habil and identified first and last names. For the most frequent first names not categorized by $gender\_guesser$, we manually assigned the gender as categorized by a native German speaker who is part of the research team.