NBER WORKING PAPER SERIES

THE EFFECT OF INFRASTRUCTURE ON WORKER MOBILITY: EVIDENCE FROM HIGH-SPEED RAIL EXPANSION IN GERMANY

Daniel F. Heuermann
Johannes F. Schmieder

Working Paper 24507
http://www.nber.org/papers/w24507

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2018

We thank Melanie Arntz, Konstantin Büchel, Colin Cameron, David Card, Wolfgang Dauth, David Dorn, Steve Gibbons, Stephan Heblich, Rick Hornbeck, Frauke Kreuter, Patricia Melo, Joachim Möller, Henry Overman, Michael Pflüger, Ninette Pilegaard, Andres Rodr guez-Pose, Martin Rosenfeld, Bastian Stockinger, Jens Südekum and Stefan Witte as well as seminar participants at the SMYE 2014, the SOLE 2014, the ERSA 2016, the Standing Field Committee for Regional Economics and in Boston, Regensburg, Dortmund, Manchester, Hamburg, Trier, Paris, Madrid, Düsseldorf and at the IAB for many helpful comments on this project. Peter Haller, Christoph Rust and Deborah Goldschmidt provided excellent research assistance. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Effect of Infrastructure on Worker Mobility: Evidence from High-Speed Rail Expansion in Germany
Daniel F. Heuermann and Johannes F. Schmieder
NBER Working Paper No. 24507
April 2018
JEL No. J61,R12,R23,R40

ABSTRACT

We use the expansion of the high-speed rail network in Germany as a natural experiment to examine the causal effect of reductions in commuting time between regions on the commuting decisions of workers and their choices regarding where to live and where to work. We exploit three key features in this setting: i) investment in high-speed rail has, in some cases dramatically, reduced travel times between regions, ii) several small towns were connected to the high-speed rail network only for political reasons, and iii) high-speed trains have left the transportation of goods unaffected. Combining novel information on train schedules and the opening of high-speed rail stations with panel data on all workers in Germany, we show that a reduction in travel time by one percent raises the number of commuters between regions by 0.25 percent. This effect is mainly driven by workers changing jobs to smaller cities while keeping their place of residence in larger ones. Our findings support the notion that benefits from infrastructure investments accrue in particular to peripheral regions, which gain access to a large pool of qualified workers with a preference for urban life. We find that the introduction of high-speed trains led to a modal shift towards rail transportation in particular on medium distances between 150 and 400 kilometers.

Daniel F. Heuermann
University of Regensburg
Department of Economics
93040 Regensburg, Germany
daniel.heuermann@ur.de

Johannes F. Schmieder
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and IZA
and also NBER
johannes@bu.edu

1 Introduction

Throughout history, infrastructure has shaped the distribution of population and economic activity by determining the cost of mobility. In line with this notion, recent empirical studies provide compelling evidence that access to road or rail transportation raises employment, wages, and GDP within regions (Gibbons, Lyytikäinen, Overman, and Sanchis-Guarner, 2012; Michaels, 2008; Banerjee, Duflo, and Qian, 2012), and positively affects the volume of interregional trade (Duranton, Morrow, and Turner, 2014). These benefits are, however, not spread evenly across space. In particular, peripheral regions have been shown to suffer losses in output and employment from firms moving to larger cities (Faber, 2014; Qin, 2014). Within cities, improvements in road infrastructure support processes of suburbanization and a corresponding decline in central city population (Baum-Snow, 2007; Garcia-López, Holl, and Viladecans-Marsal, 2015). While the effects of infrastructure on interregional trade flows and on the spatial organization of economic activity within metropolitan areas are comparatively well understood (see Redding and Turner (2015) for a survey), one aspect that has received less attention so far is that better roads and railway connections also enhance the set of options available to workers to separate their region of residence from their region of work. A natural question to ask is therefore how workers adjust their locational decisions and their commuting behavior to reductions in travel times between regions. Shedding light on this issue not only allows for gaining insight into the time elasticity of commuting decisions. It is also an opportunity to better understand the relative attractiveness of large and small cities as places of living and places of work and to examine on how benefits from infrastructure investments are distributed between urban centers and the periphery.

Identifying the causal effect that changes in travel time have on commuter numbers and on the locational choices of workers is difficult for two reasons. First, infrastructure investments are endogenous to both labor and product markets. The existing literature has frequently dealt with this issue by using either planned (Duranton and Turner, 2012; Duranton, Morrow, and Turner, 2014) or historical (Garcia-López, Holl, and Viladecans-Marsal, 2015) roads, or least-cost corridors (Faber, 2014) as instrumental variables. Second, transportation infrastructure in most cases jointly affects the mobility of workers and goods (Monte, Redding, and Rossi-Hansberg, 2015). This complicates the identification of the effect that commuting times have on locational decisions of workers due to confounding effects from interregional trade. This second issue is ideally addressed by means of an instrument that exclusively affects the mobility of workers while leaving the costs of goods transportation unchanged.[1]

In this paper, we use the expansion of the high-speed rail (HSR) network in Germany as a natural experiment to isolate the effect that changes in travel time between regions have on commuting behavior and the locational decisions of workers. This setting exhibits three features which are key for our empirical design. First, due to the

[1] While the current literature has drawn on natural experiments to shed light on the existence of agglomeration economies (e.g., Davis and Weinstein (2002) and Ahlfeldt, Redding, Sturm, and Wolf (2015)), these studies usually do not differentiate between product and labor market channels. A notable exception is Redding and Sturm (2008), who draw on the history of German division and re-unification to examine the effects of changes in market access on city growth.

large-scale nature of the investment, the introduction of high-speed trains has substantially reduced commuting times between regions. For a number of peripheral counties, which over decades were badly connected to the rail network, these time savings have been dramatic. The opening of direct HSR-connections has in some cases reduced travel time by more than 70 percent, bringing large urban labor markets into daily reach. The second feature we exploit is that the German HSR-network was introduced and extended in two major waves which are structurally very different. During the first wave, ranging from 1991 to about 1998, all large cities in Germany were successively connected to HSR. However, due to the unequal distribution of these cities across states (*Bundesländer*), the majority of the 46 HSR-stations existing in 1998 were located in only 3 out of the 16 German states (*North Rhine-Westphalia*, *Baden-Württemberg*, *Hesse*). Several states through which the tracks ran but which had obtained none or only a small number of stops therefore vehemently demanded HSR-stations to also be established on their territory. As a result, the network was in a second wave extended to also cover smaller cities located en route between metropolitan hubs. In these cases, the final location of an HSR-station was often the result of political haggles between national, regional and local authorities and influenced by petitions, court decisions, and interest groups. Consequently, a number of 'lucky' cities were in this phase endowed with an HSR-station while other comparable cities were not. A prime example is Montabaur, a small town with slightly above 12,000 inhabitants, which was connected to HSR in 2002 as part of the new high-speed connection between Frankfurt and Cologne. As shown by Ahlfeldt and Feddersen (2017), the opening of this HSR-station was exogenous to the levels and trends of local economic development. The same argument applies to virtually all counties that were endowed with an HSR-station between 1999 and 2010. Based on this idea, we use the introduction of a direct HSR-connection between two regions during this period of network expansion as an instrument for commuting time. A third feature which renders this setting well suited for identifying the effect that changes in commuting time have on the number of commuters between regions is that high-speed train connections effectively increase the size of the local labor market while leaving product markets unaffected. Since its introduction in 1991, the German high-speed train, the InterCityExpress (ICE), has been used only for passenger transportation.[2] While the idea to also use the ICE for freight transportation had initially been considered, it was discarded already in early stages of technology development. In contrast to large-scale investments in highways and roads, this exclusive focus on passenger transportation makes the expansion of the HSR-network an ideal case to isolate the effect of travel time on the commuting behavior of workers.

In order to exploit the quasi-randomness of changes in travel times between regions, we construct a novel data set on the establishment of ICE-stations, connection frequency, and travel times by train between counties, and combine this information with individual employment data for all workers in Germany. Based on these data, we estimate a gravity model with instrumental variables where we use the opening of a direct ICE-connection

[2] According to information provided by the Deutsche Bahn, 81 million passengers have used the ICE in 2013 (Deutsche Bahn, 2014).

between two regions as an instrument for commuting time. Beyond their importance for our identification approach, the results from this first stage provide novel evidence on the average effect that high-speed trains have had on commuting times between regions in Germany. We use these results to examine the effect that changes in travel time have on the number of workers commuting across county borders. We then proceed by analyzing the locational choices that workers make in the face of reductions in commuting time in greater detail. In particular, we shed light on the margin and the direction of adjustment, i.e., we examine whether changes in commuting behavior are directed towards large or small cities and whether they are the result of workers changing jobs or places of residence. These results are instructive in two respects. First, they shed light on the consumption value of urban areas as places of living. Second, they provide evidence on how benefits from large-scale infrastructure investments are distributed between different types of regions. As a final step, we examine whether reductions in travel time and in particular the introduction of direct HSR-trains have an influence on the choice that workers make regarding the mode of transportation they use on their way to work.

We find that the expansion of the HSR-network has reduced travel times on average by 9.5 percent on relations that were endowed with a direct ICE.[3] The results from the second stage suggest that a reduction in commuting time by one percent causes a rise in commuter numbers of 0.25 percent. This rising incidence of commuting is largely driven by workers who change their place of work from larger to smaller cities while retaining their place of living in urban areas. This finding supports the notion that the amenity value of cities has an influence on the locational choice of workers and, as such, is an important factor for existing commuting patterns. It also suggests that benefits from infrastructure investments accrue largely to smaller cities, which gain access to a larger pool of qualified workers. Finally, the results show that investments in rail infrastructure have substantially raised the share of rail-based commuting, in particular on relations that were endowed with a direct ICE-connection.

The next section outlines the evolution of the HSR-network in Germany and explains the rationale for using the opening of direct ICE-connections as an instrument for travel time. In Section 3, we describe the data and provide descriptive statistics on regions and rail connections. In Section 4, we explain our identification approach. Section 5 provides the results. Section 6 concludes.

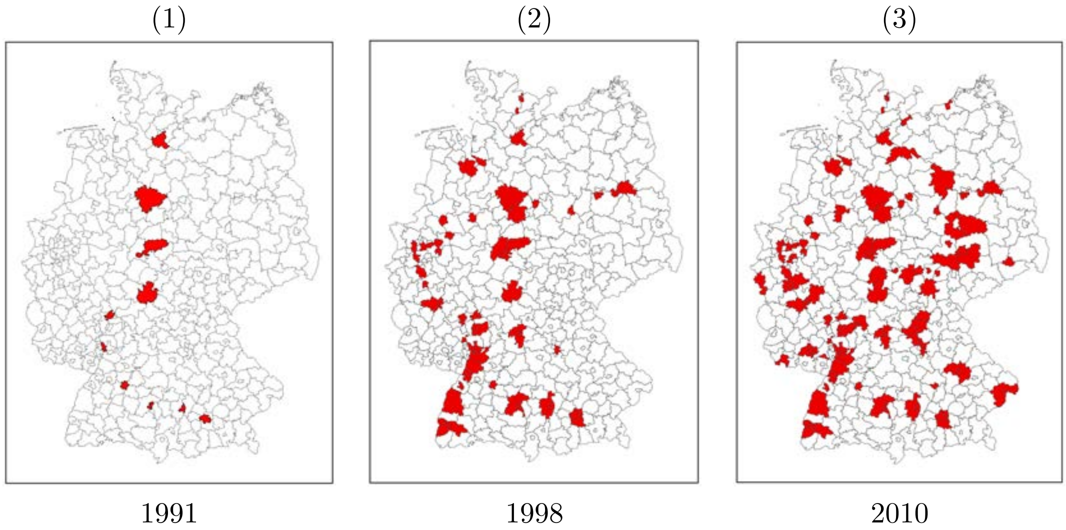
2 HSR as an Instrument for Commuting Time

2.1 Evolution of High-Speed Rail in Germany

Since the 1970s, transportation policies in most developed countries have undergone a re-focus from motorized individual transport to railway transportation, a shift sometimes referred to as the 'renaissance of the train' (Givoni, 2006). In Japan, the first high-speed

[3] Note that a *connection* between two regions can be disaggregated into two unidirectional *relations*. By means of an example, the train connection between Frankfurt and Montabaur consists of two relations, one going from Frankfurt to Montabaur and the other one from Montabaur to Frankfurt. While the opening of a direct ICE-connection affects travel time on both relations to the same degree, commuter numbers may react differently on the two relations.

Figure 1: Evolution of the High-Speed Rail Network



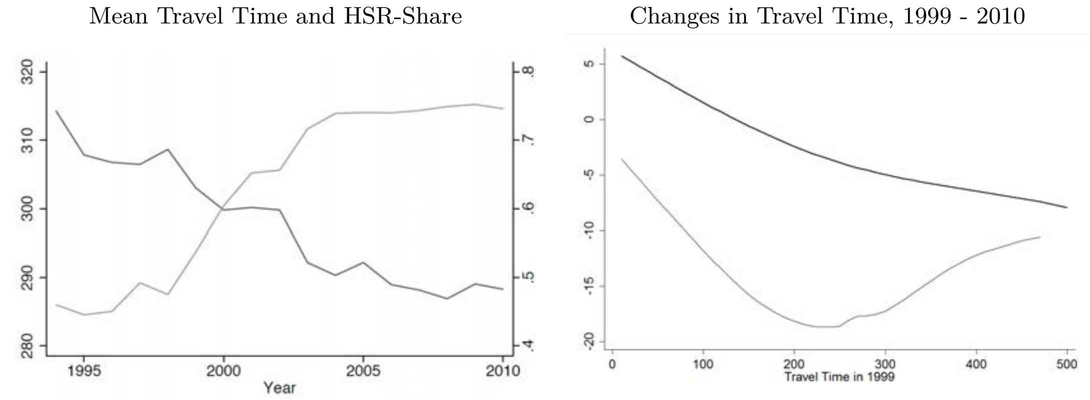
Red color indicates that a county is connected to the high-speed rail network.

train, the Shinkansen, was introduced already in 1964. Italy ('Pendolino') and France ('TGV') followed in 1977 and 1981, respectively. The HSR-network in Germany was opened in 1991 after two decades of publicly funded research into tracks and propulsion technology.[4] As shown in Figure 1, the high-speed rail era started with one major track running from Hamburg via Mannheim and Frankfurt to Munich, serving a total of ten stations. The network has since then been enlarged in two major waves. The initial north-south axis was in the following years complemented by an east-west connection from Berlin to Cologne, covering several large cities in the Rhine-Ruhr area, and a diagonal line from Cologne to Munich via Stuttgart, Augsburg, and Ulm. In addition, the branches to Kiel, Bremen and Freiburg were added to the network. In 1998, a total of 46 ICE-stops were in operation. At this point, all major cities in Germany were connected to the ICE-network and the number of direct connections amounted to 318.

The second wave started with the opening of the Berlin-Munich line in 1999, when a number of smaller cities in Eastern Germany and Bavaria, including Naumburg, Saalfeld, and Lichtenfels, were connected to the network. In 2002, the towns of Montabaur, Limburg, and Siegburg, which are located en route between Cologne and Frankfurt, gained access to the new HSR-line built between the Rhine-Main and the Rhine-Ruhr area. During this period, the network was extended to cover 34 additional small and medium-sized cities. In 2010, 80 counties were served by 260 high-speed trains in at least four hour intervals, in most cases every one or two hours. As as result of these extensions, the number of direct ICE-connections rose by 582 to a total of 900 in 2010.

[4] In contrast to most other countries, high-speed trains in Germany are for the most part running on the same tracks as other passenger trains and freight trains (referred to as *fully mixed model* by Barrón, Campos, Gagnepain, Nash, Ulied, and Vickerman (2009)). Stations and tracks were usually adjusted to the requirements of the ICE and only in a small number of cases newly built for the exclusive use of high-speed rail traffic.

Figure 2: High-Speed Rail and Travel Time



Left Panel: the black line (left scale) shows the average travel time of the fastest connection between all counties per year; the gray line (right scale) provides the share of these connections that entail the use of an ICE-train on at least one leg of the trip. Right Panel: the gray line shows the growth of average travel time in percent on all relations that were endowed with a direct ICE-connection between 1999 and 2010; the black line displays the growth of average travel time on all other relations.

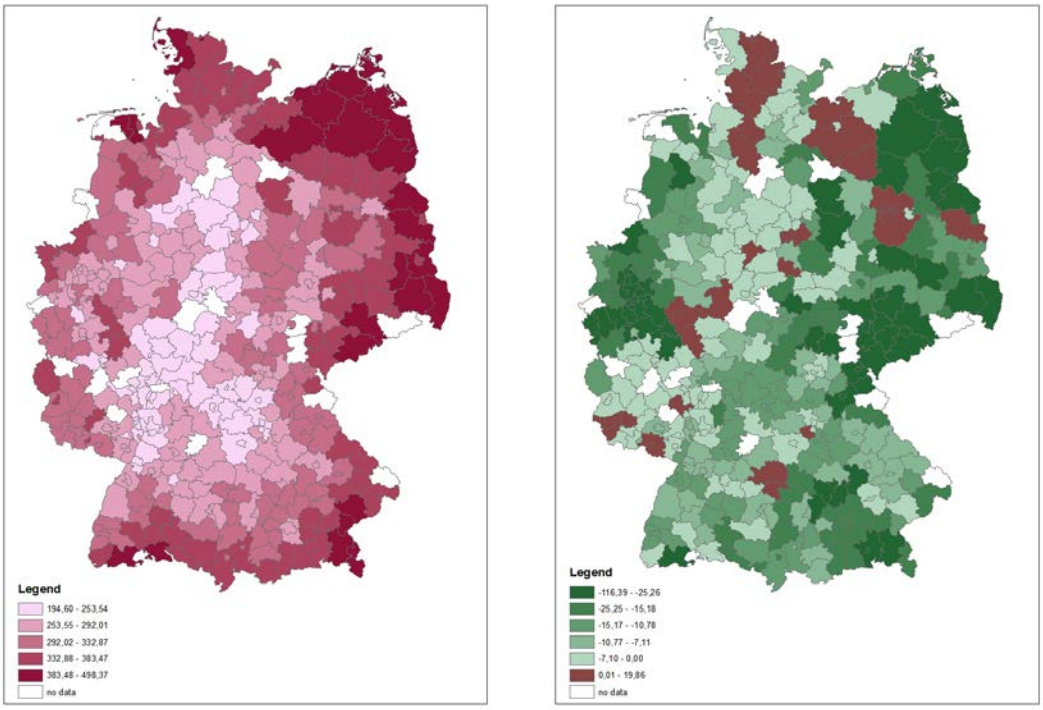
2.2 HSR and Travel Time (*Instrument Relevance*)

Our identification strategy relies on using the expansion of the high speed rail network as an instrument for travel time between regions. We therefore created a unique new data set from electronic train schedules, which we describe in more detail in the next section. In essence, this data set contains information on travel times and train connections between all regions in Germany over the period from 1994 to 2010. These data clearly show that since their introduction in 1991, high-speed trains have contributed substantially to a reduction in travel times between regions. In the left panel of Figure 2, we relate the average travel time between county capitals to the share of fastest connections that encompass the use of an ICE-train on at least one leg of the trip. Between 1994, which is the earliest year for which electronic schedule data are available, and 2010, the share of connections encompassing the use of an ICE-train has risen from 48 to 75 percent. During the same period, average travel time has decreased by nine percent from 315 to 287 minutes. Significant reductions occurred mostly at points in time when major ICE-lines were opened. Prominent cases are the opening of new stations in 1998 for the *EXPO 2000* (World's Fair) in Hanover and the introduction of the ICE-connection between Berlin and Munich between 1999 and 2000. Similarly, the nine-minute drop in average travel time between 2002 and 2003 partly goes back to the new high-speed track between Frankfurt and Cologne (see also Figure 4). The right panel illustrates how changes in travel times during the second wave of network expansion vary with initial travel time in 1999. The gray line contains average changes in percent for all relations that were endowed with a direct ICE-connection during this period. The black line provides these averages for all other relations. A comparison of both lines shows that with an average of 13 percent, reductions in travel time are particularly pronounced on relations endowed with a direct ICE-connection between 1999 and 2010. The average reduction on all other relations amounts, in contrast, to only 3.5 percent.

The first map in Figure 3 shows the average travel time needed to get from each of the

Figure 3: Average Travel Time, by Region

Average Travel Time in 1999 | Changes in Travel Time, 1999 - 2010

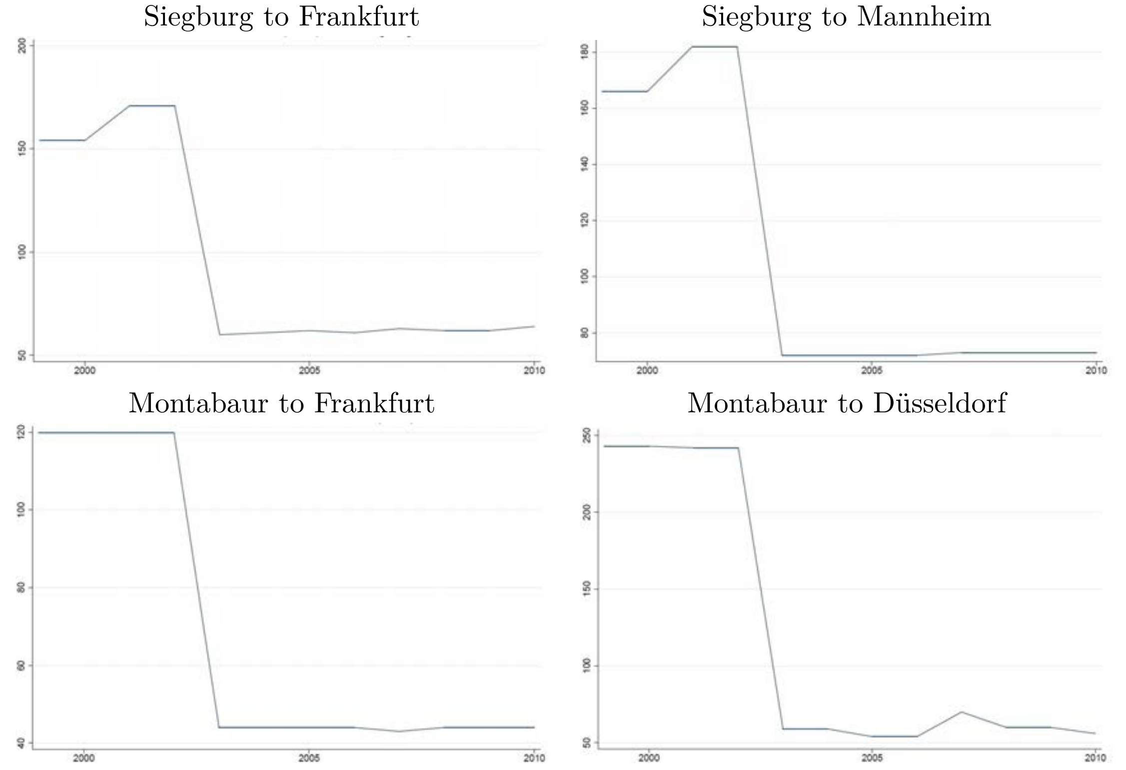


The left map shows average travel times by region in minutes for the year 1999. The right map contains changes therein in absolute values over the period from 1999 to 2010.

352 regions to any other region by train in 1999. The pattern closely follows the routes of the HSR-lines that were established during the first period of network introduction. The regions that were best connected at this point were the large cities within the corridors Hamburg - Frankfurt - Mannheim - Freiburg and Dortmund - Frankfurt - Nuremberg - Munich. Peripheral regions, in particular in the eastern part of the country, had so far remained unaffected from the introduction of HSR and exhibited substantially larger travel times. The second map contains changes in average travel times between 1999 and 2010 in absolute values. It shows that first and foremost counties in Eastern Germany, in Bavaria, and around the Rhine-Ruhr-area, which were connected to the Berlin-Munich and the newly built Cologne-Frankfurt line, saw substantial reductions in average travel time, in some cases of more than one hour.

For these counties, which over decades often had only poor access to the rail network, the opening of an HSR-station brought dramatic improvements in accessibility. The towns of Siegburg and Montabaur are prime examples. Both are located en route between Frankfurt and Cologne and were connected to the HSR-network in 2002. Figure 4 shows the average travel time by year from Siegburg to Frankfurt and to Mannheim, and from Montabaur to Frankfurt and to Düsseldorf. Travel times have dropped between 59 (Siegburg-Mannheim) and 77 percent (Montabaur-Düsseldorf) with the opening of a direct ICE-connection. Important in the present context, in each of these cases travel time has fallen from prohibitively high levels to a magnitude that can be reconciled with

Figure 4: High-Speed Rail and Travel Time - Case Study Evidence



Panels show the duration of the fastest train connection between origin and destination region in minutes.

daily commuting.[5] Overall, the evidence on both aggregate and case study level suggests that the introduction of direct HSR-connections has significantly reduced travel times between regions, lending support to its relevance as an instrument for commuting time.

2.3 Allocation of HSR-Stations (*Instrument Exogeneity*)

Most ICE-stations that were opened during the second wave of network expansion are located in small cities along ICE-connections that had already been established before 1999 with the intent to connect two larger metropolitan areas. Table A.3 in the Appendix contains a list of all regions that were endowed with an ICE-station in either the first or the second wave and compares their ranks in terms of population size.[6] With the exception of Leipzig and Dresden, the 24 largest cities in Germany were all connected to the HSR-network between 1991 and 1998. Out of the 46 cities that got access to high-speed transportation during this phase, only five are not among the largest 100 cities in Germany. The median rank in terms of city size is 28.5 and the mean population size ranges slightly above 400,000 inhabitants. The picture for the second wave is different. None of the 34 cities which were endowed with an ICE-station between 1999 and 2010 is among the largest ten and only two (Leipzig and Dresden) are among the first twenty.

[5] Road distances between these cities are 161km (Siegburg-Frankfurt), 218km (Siegburg-Mannheim), 98km (Montabaur-Frankfurt), and 135km (Montabaur-Düsseldorf).

[6] Note that the city of *Limburg an der Lahn*, which features prominently in the study by Ahlfeldt and Feddersen (2017), is not included in the list because it was not served by an ICE-train between 6 a.m. and 9 a.m. during the period of observation (see section 3.1 for details regarding the construction of the data set).

The median rank is 119 and the average population number amounts to 117,000 with 20 cities exhibiting less than 100,000 inhabitants. The reason why these regions despite their comparatively small size also got access to the HSR-network over time is that the states situated en route between larger metropolises had been obliged to contribute to the costs of upgrading the rail network for the use of ICE-trains and in return demanded the establishment of stops on their territory. The subsequent negotiations between the Federal Government, the national railway company *Deutsche Bahn AG* and the states on the location of these intermediate ICE-stations were shaped by controversy and political interest. The final choice of a location during this period was often the result of political opportunism, extensive lobbying by counties and municipalities, an intense public debate, court decisions, and topographical and geological considerations. Ahlfeldt and Feddersen (2017) provide a detailed description of the twists and turns that the political, legal, and administrative process had gone through before the ICE-stations in Siegburg, Limburg and Montabaur were finally approved and built. Importantly, future commuter numbers played virtually no role for the final assignment of an ICE-station. This disregard, which was of vital importance for small cities to become potential candidates for the opening of an ICE-stop, was made possible by the fact that prevailing legal provisions fail to consider future demand as a criterion for determining the cost-benefit ratios of large-scale infrastructure projects (Becker, 2016). In fact, the dimensions that need to be positively evaluated in order for a project to be included in the Federal Transport Infrastructure Plan (*Bundesverkehrswegeplan*) encompass its predicted effects on 1) transportation costs, 2) maintenance costs, 3) number of accidents, 4) travel time for passenger transportation, 5) travel time for freight transportation, 6) four environmental dimensions including pollutant and noise emission, 7) reliability, and 8) reductions in travel time for other means of transportation. These dimensions not only ignore future commuter numbers. In addition, the need to assign monetary values to these dimensions leaves substantial room for the states, counties and municipalities to influence the cost-benefit ratio of a project. In his analysis of the susceptibility of the Federal Transport Infrastructure Plan to political manipulation, Becker (2016, p.13) concludes that '*the result is unambiguous: a state government or any other responsible body can by means of her actions support or impede those projects that she deems suitable or not.*'[7]

The political nature of the planning process and the regulations in public procurement law not only render the regional assignment of ICE-stations during the second wave independent of future commuter numbers. The legal and administrative complexity of public procurement and the challenge of steering such large-scale projects also introduce considerable uncertainty regarding the exact timing of the opening of a direct ICE-connection. A large literature in law, including directives, legal commentaries, and handbooks on how to participate in a tender, documents the hurdles that need to be cleared before political and administrative decisions are finally taken and building permits and contract awards are legally valid.[8] As a result, in their meta-study on

[7] Own translation. Original version: '*Das Ergebnis ist eindeutig: Eine Landesregierung oder ein anderer Projektträger kann durch seine Handlungen diejenigen Projekte befördern oder hemmen, die ihm mehr oder weniger geeignet erscheinen.*'

[8] Hantschel and Schlange-Schöningen (2014) provide a summary of public procurement law and the

time delays and cost overruns in 170 large and publicly funded infrastructure projects, Kostka and Anzinger (2015) conclude that the majority of these projects were finished late and/or at significantly higher costs than planned. The study explicitly mentions the newly built HSR-line from Cologne to Frankfurt, '*which was planned to cost about €4 billion in 1995. Because of economic complexity, legal issues and problematic stakeholder relations, the project was delayed from a scheduled finish in 1999 to 2002 and increased by almost 52 percent in cost*' (Kostka and Anzinger, 2015, p.9).

Exploiting this quasi-randomness in the allocation of ICE-stations to small cities located between between major metropolitan hubs stands in the tradition of 'inconsequential places approaches' (Chandra and Thompson, 2000), which have been used among others by Michaels (2008) and Qin (2014) to identify the effect of infrastructure on local GDP, interregional trade, and the local composition of skills. Beyond their suitability as natural experiments, these relations are of particular interest in the present context because for workers in peripheral regions the opening of an HSR-station often brought large urban labor markets like, e.g., Frankfurt, Cologne, and Munich, into daily reach. Hence, in addition to an analysis of the time elasticity of commuter numbers this setting allows to shed light on the distribution of benefits from infrastructure investments between small and large cities as well as on the amenity value of urban centers.

3 Data and Descriptives

3.1 Data

For our analysis, we draw on two main data sets. First, we make use of electronic train schedules that were published on CD ROMs twice a year from 1994 until 2010 by the German railway company *Deutsche Bahn AG*. These schedules contain information on all passenger train connections between all 5,400 train stations in Germany.[9] From these CDs we extracted the travel times between all regional capitals in Germany for the years 1994 to 2010 using techniques similar to automatized web-scraping. For every year, we took the first Monday in June that is not a holiday and calculated the trip duration of the fastest connection in the time window between 6:00 a.m. and 9:00 a.m., i.e., a typical commuting time. In addition, we extracted the number of times a traveler has to change trains on this connection as well as the types of trains she would use. These latter two pieces of information allow us to determine the year of establishment of an ICE-station and to identify whether two regions are connected by a direct ICE. The frequency by which stations or relations are served by ICE-trains varies substantially. Throughout this paper, we define ICE-stations and ICE-connections as those which are served by ICE-trains in at least four-hour intervals.

We then merged the information on trip duration and ICE-connections with the universe of worker data from the administrative records of the Federal Employment

legal and administrative process of decision making. The dissertation by Vogel (2009) describes in detail the challenges encountered in the steering of public superstructure projects.

[9] The *Deutsche Bahn AG* was not willing to provide us with schedule data. We therefore resorted to the CDs and obtained an almost complete set from various book/media-sellers and via ebay.

Agency provided by the Institute for Employment Research (IAB). These data cover all employed persons subject to statutory social security contributions and contain information on education, age, gender, nationality, full-time vs. part-time employment, occupation, wages, firm size, and industry. Important for our purpose, from 1999 onwards the data set contains information on the regions that each worker lives and works in. From the full sample of workers, we generated an annual panel data set of all full-time employed workers as of June 30^{th} of each year. To this, we added information on regional population numbers and county size provided by the Federal Statistical Office (*Statistisches Bundesamt*). The geo-coordinates of each county capital are obtained from www.wikipedia.org. Based on these coordinates we calculated the pairwise Euclidian distance between all county capitals.

Throughout the paper, regions are defined on NUTS-III level. As of 2010, the 402 regions in Germany consist of 107 cities (*Kreisfreie Städte*) and 295 mostly rural counties (*Landkreise*). 32 cities are for historical reasons split up into a core city and the surrounding hinterland. In most of these cases, the city and the corresponding county carry the same name (e.g., *Munich (Stadt)* and *Munich (Land)*). Since both counties together effectively constitute the overall city area and are served by the same main train station, we have merged them into one. Out of the resulting 370 regions, 17 regional capitals are not connected to the rail network. For one further region, the electronic schedules do not contain information on train connections. Table A.1 in the Appendix lists the regions we have merged and those not contained in the data. Overall, the data set contains information on 352 regions that are accessible by railway. Theoretically, 123,552 unidirectional relations can exist between the regional capitals. For some (in particular very distant and/or rural) county pairs no meaningful train connection can be found in the data. In total, the electronic schedules contain information on travel times, train types, and the number of train changes for 1,191,733 year-relation observations over the period from 1999 to 2010. Table A.2 in the Appendix provides the number of observations per year. For each of these relations, we calculated the number of commuters from the worker data and merged it to the schedule data.

3.2 Descriptive Statistics

The upper part of Table 1 provides descriptive statistics for the 352 NUTS-III regions. The lower panel contains information on commuter numbers, distance, travel time and transported goods for the unidirectional relations between these regions.

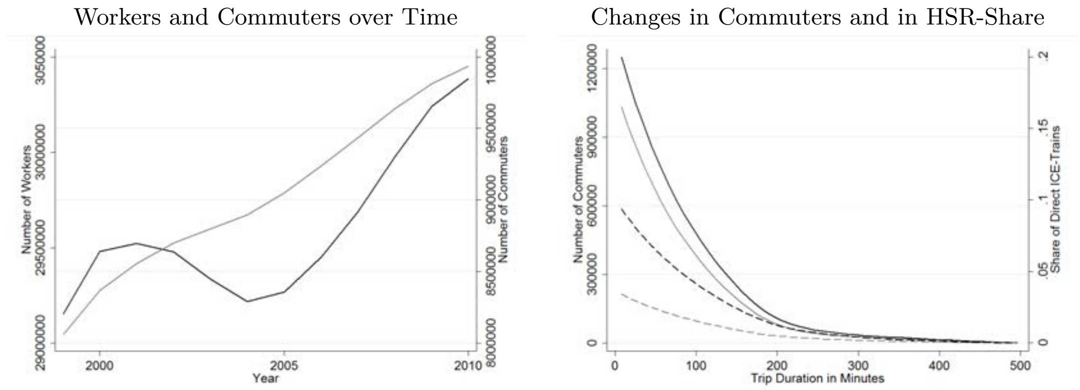
Each region covers an average area of 968 square kilometers, which is about 25 percent less than the mean size of a county in the US. The average population number is 240,200. Out of these, 90,700 persons are on average full-time employed. Due to their comparatively small size, NUTS-III regions are not self-contained labor markets. Rather, one core city is surrounded on average by five smaller and usually more rural counties (Haller and Heuermann, 2016). As a result, one third of all workers are employed outside their home region. The city of Fürth exhibits the largest share of out-commuters (73 percent), which results from its close proximity to Nuremberg as the regional capital and to the location of a large production site of *Siemens* in the nearby city of Erlangen.

Table 1: Descriptive Statistics, 1999 - 2010

	Mean	SD	Min.	Max.
Area per Region	968.4	949.1	35.6	5,811
Population per Region	240.2	276.7	33.9	3,461
Workers per Region	90.7	93.3	12.7	1,128
Out-Commuters per Region	27.4	19.1	1.9	157
Share of Out-Commuters	33.9	12.8	9.4	73
Commuters per Relation	91.0	853.2	1	57,030
Distance per Relation	360.1	184.0	2.7	990
Travel Time per Relation	277.2	119.9	2	785
Goods Transported per Relation	22.1	234.4	0	22,300

Population, workers living in a region, out-commuters, and sum of transported goods are in 1,000s. Area is measured in square kilometers, population, workers, and commuters in persons, share of out-commuters as percent of working population living within one region, distance in kilometers, travel time in minutes, and transported goods in metric tons.

Figure 5: Employment, Commuting and High-Speed Rail



Left Panel: the black line (left scale) provides the number of full-time workers contained in the data by year; the gray line (right scale) shows the number of cross-border commuters by year. Right Panel: solid lines show the number of commuters (left scale) in 1999 (gray line) and 2010 (black line) by trip duration in bands of 10 minutes; dashed lines display the share of relations served by direct ICE-trains by duration band for the year 1999 (gray line) and 2010 (black line).

The average number of commuters on each unidirectional relation amounts to 91. Since at least one person commutes on each relation, the problem of zero-entries, which has received attention in the literature on international trade (see Helpman, Melitz, and Rubinstein (2008)), does not arise here because all entries in the commuter matrix are strictly positive. The average distance on a relation is 360 kilometers with a minimum of 2.7 kilometers (Neu-Ulm to Ulm, which also exhibits the smallest travel time of 2 minutes) and a maximum 990. The average travel time by train between all region pairs amounts to 277 minutes and the average amount of goods transported to 22,100 tons.

In Figure 5, we examine the relation between employment, commuting and the expansion of the HSR-network in greater detail. To get an impression of the incidence of cross-border commuting and changes therein over time, we relate the total number of workers commuting between regions to the total number of workers contained in the data (left panel). Although the number of full-time employed workers exhibits a pronounced slump during the post-millennium recession in the early 2000s, it grows from 29.5 million to 30.5 million during the period of observation. Rising from 8 million to about 10 million, the number of commuters also follows an upward trend. Notably, it lacks the dip after the turn of the century. As a result, the share of workers commuting across county borders rises from 27 percent in 1999 to 33 percent in 2010. In the right panel, we shed light on how this rising incidence of cross-border commuting is correlated with the extension of the HSR-network. The solid lines show the number of commuters in 1999 (gray line) and 2010 (black line) in ten-minute intervals of travel time. The two lines mirror the rising incidence of commuting from the left panel and, in addition, show that commuter numbers monotonically decrease with travel time. The dashed lines display the share of relations served by direct ICE-trains in 1999 (gray line) and 2010 (black line). This share has nearly doubled between 1999 and 2010 and also falls monotonically with travel time in both years. Overall, the figure suggests that the expansion of the HSR-network is positively correlated with commuter numbers over the whole range of travel times. In the next section, we explain the identification approach we use to isolate the effect that this expansion has had on the number of commuters through its impact on travel times between regions.

4 Empirical Approach

We follow the existing literature and model the relationship between travel time and the number of commuters on a relation by means of a gravity model.[10] We define C_{ijt} as the number of commuters on the relation from region i to region j in year t. The gravity model then takes the form

$$C_{ijt} = aR_{it}^{\alpha}W_{jt}^{\beta}D_{ijt}^{\gamma}, \tag{1}$$

where R_{it} is the total number of workers living in an origin region i, W_{jt} is the total number of workers employed in a destination region j, and D_{ijt} is the travel time by rail between i and j. Taking logarithms and adding an error term yields

$$\log(C_{ijt}) = \log(a) + \alpha\log(R_{it}) + \beta\log(W_{jt}) + \gamma\log(D_{ijt}) + \varepsilon_{ijt} \tag{2}$$

Estimating equation (2) cross-sectionally would suffer from several sources of omitted variable bias as cities with shorter travel durations are likely to exhibit other economic and cultural ties that influence the intensity of commuting. To partially deal with such

[10] Matha and Wintr (2009), McArthur, Kleppe, Thorsen, and Ubøe (2013), Melo, Graham, and Noland (2011) and Persyn and Torfs (2016) apply gravity models to the case of commuting but differ in the question they address. Matha and Wintr (2009) and Persyn and Torfs (2016) focus on inter- vs. intra-national border effects using the cases of Luxemburg and Belgium; McArthur, Kleppe, Thorsen, and Ubøe (2013) examine the effect of monetary costs on commuting decisions. Closest to our paper is the contribution by Melo, Graham, and Noland (2011), who investigate regional differences in the effect that physical distance has on interregional commuting in England and Wales. Their results suggest that the availability of railway infrastructure, as measured by the number of railway stations per square kilometer, is associated with significantly larger commuting distances.

sources of bias, we include relation fixed effects r_{ij} as well as year dummies t_t:

$$\log(C_{ijt}) = \log(a) + \alpha \log(R_{it}) + \beta \log(W_{jt}) + \gamma \log(D_{ijt}) + r_{ij} + t_t + \varepsilon_{ijt} \tag{3}$$

While controlling for relation fixed effects greatly strengthens the identification, changes in travel time might still be endogenous to growth in commuter numbers. Theoretically, the direction of this bias is ambiguous as transport policies may either aim at improving services on highly frequented relations or may be focused on peripheral regions with the intent to promote an equal accessibility of regions. We address this problem by using the introduction of a direct ICE-connection between two regions during the period of network expansion as an instrument for travel time. In the first stage regression

$$\log(D_{ijt}) = b + \delta \log(R_{it}) + \theta \log(W_{jt}) + \lambda ICE_{ijt} + r_{ij} + t_t + \nu_{ijt} \tag{4}$$

travel time on each relation is estimated as a function of the total number of workers living in the origin region, the total number of workers employed in the destination region, the existence of a direct ICE-connection between regions i and j at time t, and of time and relation fixed effects. Importantly, when estimating this model with relation fixed effects for the period from 1999 to 2010, λ is identified only through ICE-connections that were established during the second wave of network expansion.

When applied to the full sample of relations, the IV-model provides an estimate of the average effect that a change in travel time has on the number of workers commuting between two regions. Theoretically, such adjustments can be the result of workers changing their jobs or their places of residence between regions. In addition, either type of relocation can be targeted towards large or small cities. Based on these considerations, we examine the locational choices that workers make in the face of a drop in commuting time in greater detail. Estimating the model separately for unidirectional relations going from small to large cities and vice versa, we first differentiate our results by direction of commute. We then define two samples of workers with either a constant place of residence or a constant place of work and shed light on whether adjustments in commuter numbers are driven by changes in jobs or in places of residence.

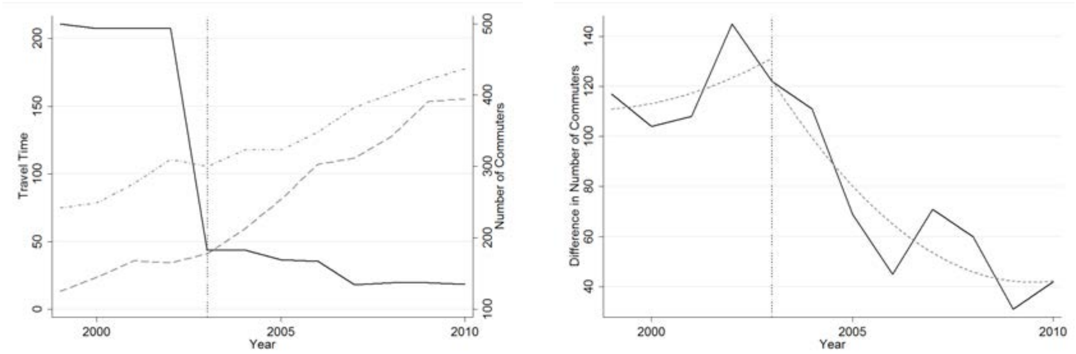
5 Results

5.1 Case Study

To gain a first understanding of the relationship between travel time and commuter numbers, we begin by examining the connection between Montabaur and Siegburg more closely. Both cities are located en route of the new railway line between Frankfurt and Cologne, which was opened in July 2002 as one key part of the Trans-European Transport Network (TEN-T) policies of the European Union (Martín and Reggiani, 2007). The objective of the project was to better connect the two largest metropolitan areas in Germany, the Rhine-Ruhr area (in the state of North Rhine-Westphalia) and the Rhine-Main area (mostly in the state of Hesse). Because the state of Rhineland-Palatinate,

Figure 6: Case Study Evidence - Montabaur and Siegburg

Travel Time and Commuters, by Direction | ∆ Commuters



Left Panel: the solid line shows the evolution of travel time between Montabaur and Siegburg (left scale); the dotted (dashed) line provides the number of commuters from Montabaur to Siegburg (Siegburg to Montabaur), both right scale. Right Panel: the solid line shows the difference in commuter numbers between both relations; dashed lines provide quadratic fits separately for the pre- and post-opening period. Vertical lines in both panels indicate the opening of the direct ICE-connection.

through which the tracks also run, demanded the establishment of a stop on its territory in return for its financial contribution, the towns of Montabaur (12,000 inhabitants) and Siegburg (41,000 inhabitants) saw themselves in the lucky position to, despite their small size, also be connected to the HSR-network. This case is close to a perfect experimental design for three reasons. First, the opening of a direct ICE-connection was unrelated to commuter numbers because both ICE-stations were built for political reasons (Ahlfeldt and Feddersen, 2017). Second, since both cities had not been connected by a direct railway line before, the drop in travel time was enormous. Third, since the new track is used exclusively by passenger trains, travel times for freight trains have remained unchanged.

The left panel in Figure 6 provides the evolution of travel time and of commuter numbers between both cities separately by direction. With the opening of the direct ICE-connection in 2002, travel time immediately fell from close to four hours to 45 minutes and from there on further to eighteen minutes in 2010.[11] Commuter numbers in both directions show an upward trend, rising from about 125 to 395 (Siegburg to Montabaur) and from 242 to 437 (Montabaur to Siegburg). While neither of the two lines exhibits a marked discontinuity in terms of commuter numbers, the adjustment process that gradually takes place on the relation from Siegburg to Montabaur after the introduction of the direct ICE-connection merits attention. Between 2003 and 2006, the number of commuters in this direction exhibits an upward shift, which moves it persistently closer to the number of workers commuting in the opposite direction. In the right panel of Figure 6, we have plotted the difference in commuter numbers between both directions by year. Before the introduction of the ICE-connection, the difference ranges between 105 and 143 commuters with a tendency to rise. This pattern changes remarkably in 2003. Over the next four years, the difference in commuter numbers falls

[11] Note that a yearly train ticket between Siegburg and Montabaur costs about 3,000 Euro per year. Doing the same trip by car takes about 50 minutes and costs about 8,000 Euro per year, assuming daily commuting.

Table 2: Direct ICE-Connections and Travel Time (*First Stage*)

Dep. Variable	(I) Dur	(II) Dur	(III) Dur	(IV) log(*Dur*)	(V) log(*Dur*)	(VI) Dur	(VII) log(*Dur*)
Direct ICE	141.3	-48.33	-12.79	-0.592	-0.095	-14.48	-0.114
	(2.639)**	(1.829)**	(0.728)**	(0.009)**	(0.006)**	(1.290)**	(0.011)**
$\log(R_{Orig})$		-25.18	29.59	0.054	0.092	10.85	-0.011
		(0.211)**	(1.256)**	(0.001)**	(0.004)**	(13.01)	(0.088)
$\log(W_{Dest})$		-25.17	20.37	0.027	0.076	-12.36	-0.001
		(0.207)**	(1.013)**	(0.001)**	(0.004)**	(9.03)	(0.051)
log(*Distance*)		147.1		0.794			
		(0.382)**		(0.002)**			
N	1,194,733	1,194,733	1,194,733	1,194,733	1,194,733	12,506	12,506
Mean Dep. Var.	277.2	277.2	277.2	277.2	277.2	163.6	163.6
Year FE	N	Y	Y	Y	Y	Y	Y
Relation FE	N	N	Y	N	Y	Y	Y
PSM	N	N	N	N	N	Y	Y

* $p < 0.05$, ** $p < 0.01$; clustered standard errors in parentheses; cluster correction on relation-level; the table shows the results from estimating equation (4) in different specifications; dependent variables are Trip Duration (*Dur*) and the log thereof (log(*Dur*)); the sample used in columns (VI) and (VII) contains only treated and untreated relations with the same ex ante probability to obtain a direct ICE-connection.

by more than 50 percent towards an average of 45 between 2006 and 2010. This finding is instructive in two respects. First, it provides anecdotal evidence that adjustments in commuting behavior after infrastructure investments fully materialize only over a period of several years. Second, it suggests that the effect of reductions in travel time on commuter numbers is larger on relations going from larger (Siegburg) to smaller cities (Montabaur) than vice versa. Importantly, this finding is closely in line with our results from the gravity model when we differentiate the effects by direction of commute.

5.2 HSR and Travel Time between Regions (*First Stage*)

Table 2 contains the results from the first stage regression. In columns (I) to (III), we specify the dependent variable in minutes to get a first impression of the time effect of a direct ICE-connection in levels. In columns (IV) and (V), we estimate the first stage defined in equation (4) with the dependent variable in logs. In the last two columns, we restrict the sample to pairs of treated and untreated relations that exhibit the same ex-ante probability of being endowed with a direct ICE-connection. We discuss these latter results in the robustness section below.

Column (I) shows that unconditional on any covariates the average travel time between two cities connected by a direct ICE is 141 minutes longer than between two cities not directly connected. This is a result of high-speed trains serving by the very nature of their technology predominantly medium- and long-distance relations. In column (II), we add the number of workers in the origin and destination region, geographic distance and year dummies to the model. The coefficient on *Direct ICE* turns negative, indicating that travel time on relations with a direct ICE-connection is on average 48 minutes shorter than on relations without. In column (III), we include relation fixed effects. As a result, γ now measures the time-effect of only those ICE-connections that were opened between 1999 and 2010. The distance measure is dropped because it is time-invariant.

The result suggests that the introduction of a direct HSR-connection within this period has reduced travel times on average by 13 minutes. Finding the time effect to be substantially smaller in the second wave compared to the first one is in line with the notion that distances between ICE-stations get smaller as the network gets denser.

Columns (IV) and (V) contain the estimates when using the log of travel duration as outcome variable. Without relation fixed effects, the introduction of a direct ICE is associated with a drop in commuting times by nearly sixty percent. In column (V), we consider only the variance from direct ICE connections established in the second wave by including relation fixed effects. The estimate shows that these newly-introduced connections have reduced traveling times on average by 9.5 percent.

Overall, the results provide evidence that the expansion of the HSR-network in the second wave has substantially reduced commuting times between regions. In line with the descriptive evidence in Section 2.2, the fact that the coefficient on *Direct ICE* is highly significant throughout all specifications suggests that the opening of a direct ICE-connection is a relevant instrument for travel time. The question is whether and to which extent workers have responded to these reductions in travel time in terms of their commuting behavior. We address this issue in the next section.

5.3 Travel Time and Commuter Numbers (*Second Stage*)

We begin by estimating equation (3) without instrumenting for travel time. The results, which are contained in column (I) of Table 3, provide an elasticity between commuter numbers and trip duration of -0.04. As discussed in Section 4, this estimate would be biased if travel time is endogenous to changes in the number of commuters, e.g., because investments in rail infrastructure are concentrated either on relations with small expected growth rates in commuter numbers (bias towards zero), or on relations with high expected growth rates (bias towards larger negative effects).

In column (II), we use the opening of a direct ICE-connection as an instrument for travel time, drawing on the specification in column (V) of Table 2 as first stage. The estimate of γ rises in absolute value, suggesting that a reduction in travel time by one percent is followed by a rise in commuter numbers of 0.25 percent. The size of the F-statistic in the first stage indicates that the introduction of a direct ICE-connection is a relevant and very strong instrument for commuting time. A comparison of both results shows that the effect of travel time on commuter numbers is substantially underestimated if the endogeneity of travel time is not controlled for. The direction of this bias is consistent with the emphasis the Federal Government has put on promoting the accessibility of peripheral regions in Eastern Germany since reunification by means of the multi-billion German Unification Transport Project (*Verkehrsprojekt Deutsche Einheit*).

The effect of changes in travel time on the number of commuters is likely to be heterogeneous with regard to the distance between home and working region for two reasons. First, reductions in travel times between proximate regions where commuting times were already small in the first place are unlikely to have a significant additional impact on the mobility of workers. Second, reductions in commuting times need to be non-negligible in order to alter the decisions of workers regarding where to live and

Table 3: Travel Time and Commuter Numbers (*Second Stage*)

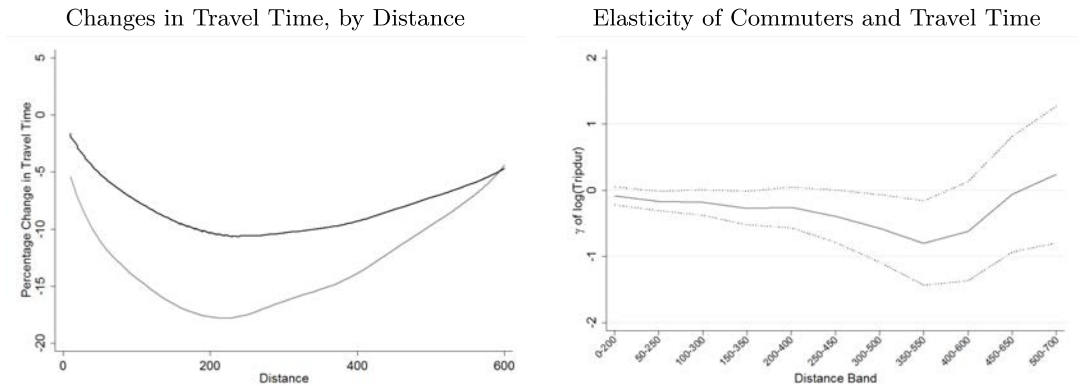
	(I)	(II)	(III)	(IV)	(V)	(VI)
Dep. Variable	log(*Number of Commuters*)					
log(*Duration*)	-0.038	-0.249	-0.239	-0.276	-0.190	-0.194
	(0.009)**	(0.086)**	(0.098)**	(0.111)**	(0.084)*	(0.094)*
log(R_{Orig})	0.255	0.277	0.272	0.314	0.337	0.664
	(0.027)**	(0.028)**	(0.031)**	(0.195)**	(0.033)**	(0.037)**
log(W_{Dest})	1.280	1.296	1.349	1.021	1.318	1.516
	(0.024)**	(0.024)**	(0.027)**	(0.183)**	(0.030)**	(0.029)**
N	1,194,733	1,191,761	1,075,396	12,264	821,808	1,191,761
Mean Dep. Var.	90.9	90.9	89.8	837.7	128.8	90.9
Year FE	Y	Y	Y	Y	Y	Y
Relation FE	Y	Y	Y	Y	Y	Y
Orig/Dest-State*Year FE	N	N	N	N	N	Y
IV	N	Y	Y	Y	Y	Y
F-Stat	-	241.3	143.8	110.5	222.6	203.0
Sample	All	All	rk>100	PSM	balanced	All

* $p < 0.05$, ** $p < 0.01$; clustered standard errors in parentheses; cluster correction on relation-level; results are those from the second stage regression in equation (3); in column (III), all relations that were endowed with a direct ICE-connection between 1999 and 2010 and where the rank of either the origin or the destination county is smaller than 100 in terms of population are excluded; column (IV) contains only pairs of treated and untreated relations with the same ex ante probability to obtain a direct ICE-connection; column (V) contains only relations with a full set of 12 observations; in column (VI), we control for a complete set of year fixed effects interacted with the state of origin (the *Bundesland* where the origin county is located in) and the state of destination.

where to work. In the left panel of Figure 7, we have plotted the average change in commuting time in percent against commuting distance in ten-kilometer intervals. On all relations combined, these reductions follow a U-shaped pattern and are largest in the range between 200 and 350 kilometers (black line). Reductions on relations that were endowed with a direct ICE-connection follow the same pattern but are about twice as large in size (gray line). With these considerations and insights in mind, one would plausibly expect the local effect of reductions in travel time to increase with distance and to peak somewhere in the middle range of the distance distribution. This is precisely what we find when estimating the IV-model separately for distance intervals of 200 kilometers. These intervals, which are shifted in steps of 50 kilometers, are chosen so as to minimize the range of the intervals while at the same time obtaining the statistical power needed to get reasonably precise results. The right panel of Figure 7 provides the point estimates and the corresponding 95 percent confidence intervals. For distances up to a range between 350 and 550 kilometers, the elasticity increases in absolute value from about -0.05 to -0.8 and again decreases thereafter. This pattern is instructive because it shows that the effect of reductions in travel times by train is most pronounced for the group of long-distance commuters. According to Pütz (2015), the number of workers commuting more than 150 kilometer to work has in recent years increased by 20 percent to 1.2 million. In Section 5.5, we discuss in greater detail to which extent reductions in travel time and the introduction of HSR-connections have contributed to this trend.

In the remaining columns of Table 3, we conduct a number of robustness checks. First, we address the concern that the regional assignment of ICE-stations during the second wave of network expansion may be correlated with expectations regarding future commuter numbers. The fact that between 1991 and 1998 all major cities in Germany

Figure 7: Heterogeneous Treatment Effects - Estimates by Distance Band



The left panel shows reductions in travel time in percent on all relations (black line) and on relations that were endowed with a direct ICE-connections between 1999 and 2010 (gray line) by distance in 10-kilometer intervals. The right panel shows points estimates and 95 percent confidence intervals for γ in bands of 200 kilometers.

were endowed with access to the HSR-network provides support to the notion that the potential of a region to attract an ICE-station rises with its economic significance. In the second wave of network expansion, this applies in particular to Leipzig and Dresden, which were by far the largest cities to gain access to the HSR-network during this period. In column (III), we therefore drop all relations from the sample which were endowed with a direct ICE-connection between 1999 and 2010 and where the rank of the origin or the destination county is smaller than 100 in terms of population. This applies to 16 out of the 34 cities which have received an ICE-station between 1999 and 2010 (see Table A.3). Dropping these relations reduces the sample by 116,365 observations. The estimate of γ falls only slightly from -0.249 to -0.239. This finding is of prime importance because it shows that the effect of reductions in travel time on commuter numbers also holds for ICE-connections between smaller cities which lack the significance to gain access to the HSR-network for economic reasons.

In a second robustness check we address the issue that less than one percent of all relations were over time endowed with a direct ICE-connection. In order to corroborate that our results are not driven by the inclusion of untreated relations which are structurally very different from the treated ones, we use Propensity Score Matching to identify treated and untreated pairs of relations with the same ex-ante probability to obtain a direct ICE-connection during the second wave (Heckman, Ichimura, and Todd, 1998). We first conduct a nearest neighbor matching without replacement, where we estimate the probability of a relation to be treated during the period of observation as a function of its characteristics in 1999. The dependent variable in the matching equation is the introduction of a direct ICE-connection on a relation between 1999 and 2010. As predictors for receiving treatment, we use the number of workers, population size, and population density in the origin and the destination county, as well as distance, travel time, commuter numbers, and the existence of an indirect ICE-connection between both counties in 1999. For both origin and destination county, we include dummy variables for the state (*Bundesland*) each county is located in and for the degree of agglomeration of the wider region according to the classification of the Federal Institute for Research on Building, Urban Affairs and Spatial Development (*Bundesinstitut für Bau-, Stadt- und Raumforschung, BBSR*). Table A.4 in the Appendix contains the results from the matching regression. It shows that most of these variables are relevant predictors for the later opening of a direct ICE-connection. For 511 of the 582 treated relations we are able to identify a nearest neighbor which has the same ex-ante probability of receiving an ICE-connection over the next eleven years. A comparison of the variable means shows that treated and untreated observations in the matched sample do not differ in terms of observables. Based on the 1,022 matched relations, we re-estimate equations (3) and (4) for the years 1999 to 2010. Columns (VI) and (VII) in Table 2 contain the results from the first stage. The mean of travel time as dependent variable falls by more than 40 percent, suggesting that the Propensity Score Matching has eliminated very distant county pairs which were ex ante very unlikely to be connected by a direct high-speed train. The time effect of a direct ICE-connection slightly increases to 14 minutes, which is equivalent to a reduction in travel time of eleven percent. The results from the second stage are contained in column (IV) in Table 3. Consistent with the first stage, the mean of commuter numbers as the dependent variable rises substantially. The effect of travel time on commuter numbers increases, however, only slightly from -0.249 to -0.276, confirming that the findings are robust to the use of a restricted control group.

Third, we address the issue that for some relations the schedule data do not contain information on travel times for all years (see Table A.2 in the Appendix). This applies in particular to earlier years and very distant relations. We therefore re-estimate the IV-model based on the 68,484 relations for which a full set of 12 observations is available. Finding the coefficient on trip duration in column (V) to fall only slightly to -0.190 suggests that the findings obtained so far also hold for a balanced sample of relations.

Fourth, we control in a more flexible way for regional economic shocks that might be correlated with the expansion of the ICE-network. If, for instance, ICE-stations were predominantly opened in fast-growing regions, the effect of the rise in economic activity on commuting might bias our results regarding the causal effect of travel time. We therefore interact fixed effects for each of the 16 German states (*Bundesländer*) of both the origin and destination counties with year effects and add these as controls in column (VI). The estimated elasticity of commuter numbers with respect to travel duration is now estimated at -0.194, which is close to the estimates from the other specifications.

The coefficients on worker numbers in the origin and destination region warrant some comments. Both remain relatively stable across all specifications. Increasing the number of workers in the origin county by one percent raises the number of commuters by about 0.3 to 0.6 percent. Increasing the number of workers in the destination county raises the number of commuters by between 1 and 1.5 percent. This yields two insights. First, the relative magnitude of the coefficients pinpoints the importance of the demand side for commuting as compared to the supply side. While the availability of jobs exerts a strong pull for workers into these regions, the supply of more workers within a region does only to a lesser extent push them into jobs in other regions. Second, the fact that both coefficients add up to more than 1 indicates that the number of commuters rises more than proportionally with the size of the joint labor market.

As a final robustness check, we examine whether our results are influenced by a confounding effect of HSR on freight transportation. In general, microeconomic adjustments to infrastructure investments are usually not restricted to the labor market. Duranton, Morrow, and Turner (2014) and Donaldson (2016) show that the volume of inter-regional trade and, in consequence, the degree of sectoral specialization within regions rises with the availability of road or rail infrastructure. Such shifts in industrial composition will lead to more intense commuting if workers change jobs across regions for a better match. Hence, if infrastructure investments also affect the volume of transported goods, this is likely to bias any estimate of the causal effect of travel time on commuter numbers.[12]

In the present context, direct effects of HSR on the volume of rail cargo are unlikely because high-speed trains are exclusively used for passenger transportation. Positive indirect effects may, however, arise if improved tracks also allow for a higher frequency of freight trains. On the other hand, congestion in the network due to a larger number of trains my negatively affect transportation volumes. Both mechanisms are particularly relevant in the German mixed-use network. We address this concern empirically by estimating the gravity model with the amount of goods transported by rail on each relation per year in metric tons (in logs) as dependent variable. Data are provided by the Federal Statistical Office (*Statistisches Bundesamt*) and are differentiated by twenty groups of goods (*Gütergruppe nach NST-2007*). As these data are available only on the level of the 39 NUTS-II regions in Germany, we aggregate the number of workers living and working in each region to NUTS-II level. The variable *Direct ICE* is set to one if at least one direct HSR-connection exists between two NUTS-II regions. In columns (I) to (III) of Table 4, we use the total amount of rail cargo transported on a relation in a given year as dependent variable. The first two columns contain the results from a reduced form of the IV-model. In column (I), the existence of a direct ICE-connection is correlated with a larger volume of transported goods. This effect vanishes, however, with the inclusion of relation fixed effects in column (II). Consistently, the results from the second stage of the IV-model, which are contained in column (III), provide no evidence for bilateral trade flows to react in any direction to changes in commuting time. In columns (IV) to (VI), we disaggregate the amount of transported goods by goods groups, i.e., one observation exists for each group of goods per relation per year. Again, neither in the fixed effects model nor in the IV-model we find evidence that the tonnage of cargo transported by rail reacts in any direction to the opening of a direct HSR-connection.

5.4 Direction of Commute and Margin of Adjustment

The results obtained so far suggest that reductions in travel time raise the number of commuters between regions by altering the choices that workers make regarding an optimal combination of home and work region. In this section, we analyze the locational decisions that workers make in the face of falling commuting times in greater detail. Two dimensions are of particular interest. The first one pertains to the direction of commute:

[12] Alternatively, firms may alter their locational decisions as a result of changes in regional accessibility. We have addressed this issue in the working paper version and have found no evidence in this regard.

Table 4: High-Speed Rail and Goods Transportation

	(I)	(II)	(III)	(IV)	(V)	(VI)
Dep. Variable	log(*Transported Goods in Tons*)					
Direct ICE	0.614	0.090		0.089	0.024	
	(0.082)**	(0.055)		(0.045)*	(0.036)	
log(*Duration*)			-3.691			-1.357
			(2.295)			(2.057)
log(R_{Orig})	0.069	-0.326	-0.238	-0.051	-0.153	-0.126
	(0.031)*	(0.065)**	(0.089)**	(0.021)*	(0.047)**	(0.065)*
log(W_{Dest})	0.190	-0.257	-0.176	0.021	-0.122	-0.098
	(0.030)**	(0.066)**	(0.086)*	(0.020)	(0.049)**	(0.064)
log(*Distance*)	-0.784			-0.318		
	(0.055)**			(0.039)**		
N	52,030	52,030	52,030	133,451	133,451	133,451
Year FE	Y	Y	Y	Y	Y	Y
Relation FE	N	Y	Y	N	Y	Y
IV	N	N	Y	N	N	Y
F-Stat	-	-	58.1	-	-	22.7
Differentiation	Sum of Transported Goods			Pooled by Groups		

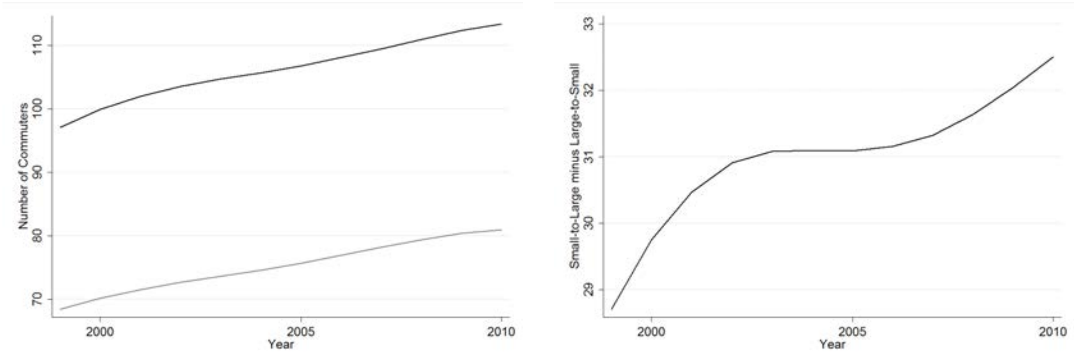
* $p < 0.05$, ** $p < 0.01$; clustered standard errors in parentheses; cluster correction on relation-level; dependent variable in columns (I) to (III) is the total amount of goods transported on a relation per year; in columns (IV) to (VI), goods are differentiated by NST-2007 classification; the model in columns (I), (II), (IV) and (V) is a reduced form of the two-stage IV-model in equations (3) and (4); columns (III) and (VI) contain the second stage of the IV-model.

Do workers start to commute from larger to smaller cities, or vice versa? The second one relates to the margin of adjustment: Do workers change jobs or places of residence as a reaction to reductions in commuting time? Providing answers to both questions is instructive in two respects. First, it allows for gaining insight into the preferences that workers have regarding their place of residence. While existing evidence suggests that improved road accessibility supports processes of suburbanization within cities (Baum-Snow, 2007; Garcia-López, Holl, and Viladecans-Marsal, 2015), less is known about how workers choose their broader region of residence in the face of a larger availability of commutable jobs. Generally, the recent literature has emphasized the changing role of cities from places of production to places of living. A rising valuation of the amenity and consumption value of urban life is often cited as one major reason for such a shift (Glaeser, Kolko, and Saiz, 2001). In addition, large cities may help to overcome the co-location problem of highly qualified and highly specialized spouses (Costa and Kahn, 2000). In light of this literature, investments into infrastructure may well support the emergence of 'hub-and-spoke' commuting relations where workers live in the city while working in the surrounding periphery.

Second, the way in which workers adjust their locational decisions to changes in travel time is likely to yield differential benefits to regions. If workers predominantly change jobs between regions due to an improved local accessibility, firms in the target regions may benefit from a larger availability of qualified workers. Alternatively, if workers first and foremost relocate their place of residence, this may alter the distribution of purchasing power across regions. The fact that travel times vary on the level of relations rather than on worker level precludes estimating the causal effect that changes in commuting time have on individual wages and consumption because it inhibits to properly control for the selection of workers into commuting. However, examining on

Figure 8: Number of Commuters by Direction of Commute

Number of Commuters, by Direction | Δ Number of Commuters



The left panel shows the average number of commuters separately for relations where the destination region is larger in terms of population than the origin region (black line) and vice versa (gray line). The right panel shows the difference between both lines in absolute terms.

which margin and in which direction workers adjust their locational decisions allows to shed light on how changes in travel time affect the distribution of jobs and population between urban areas and the periphery and to thereby indirectly gain insight into how benefits from infrastructure investments are distributed between both types of regions.

In order to address these question, we examine in a first step whether commuting intensifies from large to small cities or vice versa. We therefore define two types of relations. Large-to-small relations are those where the population number in the origin region ('region of residence') exceeds that in the destination region ('region of work'). In contrast, on small-to-large relations the origin region is smaller in terms of population than the destination region. The left panel in Figure 8 provides the average number of commuters on each relation separately by relation type. While both lines rise monotonically over time, the absolute number of commuters is about 40 percent larger on relations running from small to large cities compared to the opposite direction. The right panel shows the difference between both lines in absolute terms. While this difference rises from 29 to about 33 commuters between 1999 and 2010, it exhibits a saddle point in the middle of the period, which coincides with the opening of the major ICE-line from Frankfurt to Cologne. The question is whether reductions in commuting time by train have contributed to this temporary reversal in commuting trends. In order to address this question, we estimate the IV-model separately for both groups of relations.

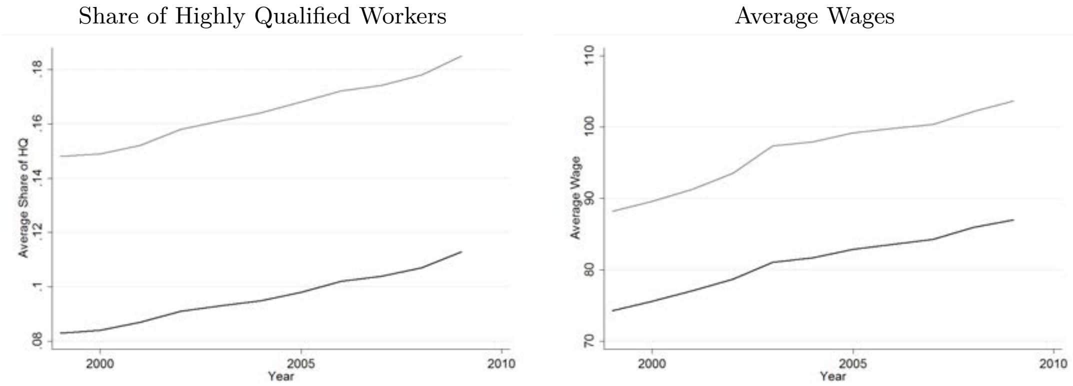
The first three columns in Table 5 provide the results from the second stage where we include only small-to-large relations. Conversely, columns (IV) to (VI) contain only large-to-small relations. Within each of the two blocks, we vary the minimum size difference between origin and destination region from 0 to 100,000 inhabitants. A comparison of both sets of results shows that the effect of reductions in commuting times on commuter numbers is exclusively driven by workers who start to commute from large to small cities. This effect varies little with size difference. Finding the number of commuters to rise first and foremost on relations from large to small cities can theoretically be the result of two types of adjustments. First, workers living in urban areas may take up jobs in the periphery which have become more easily accessible as a result of improved

Table 5: Direction of Commute and Margin of Adjustment

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
	Region of Work > Region of Residence				*Region of Work < Region of Residence*			
Dep. Variable				log(*Number of Commuters*)				
log(*Duration*)	-0.139	-0.142	-0.127	-0.381	-0.367	-0.405	-0.267	-0.134
	(0.119)	(0.121)	(0.148)	(0.126)**	(0.129)**	(0.150)**	(0.152)*	(0.100)
$\log(R_{Orig})$	0.263	0.263	0.214	0.289	0.319	0.413	0.307	0.493
	(0.036)**	(0.037)**	(0.049)**	(0.047)**	(0.049)**	(0.065)**	(0.065)**	(0.046)**
$\log(W_{Dest})$	1.344	1.382	1.501	1.263	1.269	1.339	0.544	-0.316
	(0.038)**	(0.039)**	(0.055)**	(0.032)**	(0.033)**	(0.043)**	(0.045)**	(0.032)**
N	597,773	562,419	318,726	592,680	557,644	314,903	237,745	185,109
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Relation FE	Y	Y	Y	Y	Y	Y	Y	Y
IV	Y	Y	Y	Y	Y	Y	Y	Y
Min. Difference	1	10,000	100,000	1	10,000	100,000	1	1

* $p < 0.05$, ** $p < 0.01$; clustered standard errors in parentheses; cluster correction on relation-level; the table shows the results from the second stage in equation (3); in columns (I) to (III), the region of work is larger than the region of residence; in columns (IV) to (VIII), the opposite applies; the bottom row shows the minimum difference in population numbers applied; the last two columns contain only workers with either a constant region of residence (VII), or of work (VIII).

Figure 9: Education and Wages of Commuters and Non-Commuters



The two panels contain the share of highly qualified workers (left panel) and average wages (right panel) separately for commuters (gray lines) and non-commuters (black lines).

rail connections. Alternatively, workers might keep their jobs in the periphery and move their place of residence to larger cities. In order to examine on which margin workers adjust their locational choice to reductions in commuting time, we define two subsamples of workers. The first one consist of workers who do not change their region of residence between 1999 and 2010. For these workers, changes in commuting behavior can only be the result of job changes across regions. The second subsample contains all workers who do not change their region of work during this period. For these workers, adjustments in commuting behavior can arise only through changes in the place of residence. We estimate the gravity model separately for both samples. The results are contained in columns (VII) and (VIII) in Table 5. The coefficient of commuting time is significant for workers with a constant place of residence but not for those with a constant place of work, suggesting that workers optimize their locational decisions predominantly on the margin of employment. These results render support to the notion that workers prefer to live in larger cities and respond to reductions in travel time by extending their job search to smaller regions. As such, the introduction of direct HSR-connections allows workers to enjoy the amenities of urban life (Glaeser, Kolko, and Saiz, 2001) and at the same time to profit from efficient labor market matches in the periphery. One example, which repeatedly features in the German media, is the early-morning ICE from Berlin to Wolfsburg, which covers the 200 kilometers in slightly more than one hour and is used daily by 210 employees working at *Volkswagen* (Henneke, 2011).

These findings suggest that benefits from infrastructure investments accrue mainly to smaller cities, which gain access to a larger pool of qualified workers. In Figure 9, we compare commuters to non-commuters in terms of education and wages. Commuters exhibit nearly twice the share of college graduates and earn about 20 percent higher wages compared to non-commuters. By gaining access to larger labor markets through improvements in commuting infrastructure, peripheral regions are likely to benefit from this positive selection of workers into commuting. This notion is in line with Ahlfeldt and Feddersen (2017), who argue that the rise in GDP of 8.5 percent incurred by regions as a result of being connected to the HSR-network is partly caused by gains from labor pooling and knowledge spillovers. This result is particularly important in the German context, where the headquarters or main production sites of several large and highly-productive firms are located in smaller cities rather than in the urban centers. Prominent examples are *Audi* (located in Ingolstadt), *Volkswagen* (Wolfsburg), *Siemens* (Erlangen), and *Jenoptik* (Jena). Beyond these prominent firms, numerous 'hidden champions' which sell their products on world markets are for historical reasons located in smaller towns. Substantial reductions in travel time put these firms into daily reach for highly qualified workers who prefer to live in larger cities but possess qualifications sought for by technology-intensive firms.

5.5 Rail Infrastructure and Commuting: Assessing the Evidence

How large is the effect that reductions in travel times by rail have on the number of commuters in absolute terms? Calculating the magnitude of this effect not only furthers our understanding of the extent to which faster railway connections have contributed to the rising incidence of commuting between regions in Germany. Relating it to changes in the total number of commuters also provides an estimate of the influence that reductions in travel times have on the modal choice of workers between automotive mobility and rail transportation. This diversion of traffic from road to rail is of particular interest from the perspective of environmental and transport policies as it has been shown to reduce road traffic externalities like fatalities and pollutant emission (Lalive, Luechinger, and Schmutzler, 2013). In order to address these issues, we draw on the results from the IV-model and predict the number of workers who have started to commute between regions by train as a result of reductions in travel time. Relating this number to the rise in commuter numbers between 1999 and 2010 provides an estimate of the effect that reductions in commuting time have had on the share of rail commuters.

Table 6 contains the results. Between 1999 and 2010, the total number of workers commuting across county borders has risen from close to 8 million to nearly 10 million. Given an elasticity between travel time and the number of commuters of -0.25, the average reduction in travel time of 4.6 percent during this period has caused a rise in commuter numbers of about 91,500. This implies that 4.56 percent of the additional 2 million commuters have started to commute by train *because of* reductions in travel time.[13] A natural benchmark for comparison is the share of train commuters among all 10 million commuters in Germany. According to representative information from the German Micro Census provided by the Federal Statistical Office (2016), this share amounts to 4.6 percent. This number is perfectly in line with the share predicted from our IV-model. Since our estimate does, however, not include workers who start to commute by rail for other reasons, e.g., because they do not own a car or use the train out of environmental considerations, it is very likely that the total share of rail commuters among the 2 million additional commuters exceeds 4.6 percent, which in turn suggests that reductions in travel time have led to a modal shift towards rail transportation during

[13] Note that the average reduction in travel time and the share of rail commuters are similar by coincidence, not by construction.

Table 6: Commuter Numbers and Rail Share

	(I) Commuters 1999	(II) Commuters 2010	(III) Change 1999-2010	(IV) Due to Time Reductions	(V) Additional Rail Share
			All Relations		
All Relations	7,949,000	9,952,000	2,003,000	91,414	4.56
Direct ICE	109,000	141,000	32,000	3,976	12.43
			Large to Small		
All Relations	3,259,000	4,099,000	840,000	56,967	6.78
Direct ICE	66,000	89,000	23,000	2,426	10.7

Columns (I) to (III) contain the number of commuters in 1999 and 2010 as well as the change therein in absolute terms (rounded to full 1,000s); column (IV) shows the predicted number of workers who start to commute as a result of reduced travel times by train; column (V) contains an estimate of the share of rail commuters among all additional commuters that is caused by reductions in travel time; *Direct ICE* refers to all relations endowed with a direct ICE between 1999 and 2010.

the period of observation.

In the second row, we assess the effect that high-speed trains as one particular mode of rail transportation have had on the absolute number of commuters and on the modal choice of workers. The first three columns show that the total number of commuters on the 582 relations that were endowed with a direct ICE-train between 1999 and 2010 has risen by 32,000 to 141,000. According to the results from the first stage, the introduction of HSR has led to an average decrease in travel time of 9.6 percent on relations endowed with a direct ICE-train. As shown in column (IV), these reductions were followed by a rise in commuter numbers of close to 4,000, suggesting that a share of 12.4 percent of rail commuters is caused by reductions in travel time. Notably, this share is about three times larger than the average share of rail commuters of 4.6 percent.

In the lower part of Table 6, we conduct the same calculation for relations running from large to small cities. On these relations, the number of commuters has risen from about 3.3 to 4.1 million. With an estimated elasticity between travel time and commuter numbers of 0.38, slightly more than 52,000 of these additional commuters are the result of reductions in travel time by train. The share of rail commuters that is caused by reductions in travel time amounts to 6.78 percent, which is about 50 percent larger than the corresponding share on all relations. The number of workers commuting from large to small cities on relations endowed with a direct ICE-connection has risen from 66,000 to 89,000. The opening of direct ICE-connections has contributed about 2,500 to the additional 23,000 commuters, implying that 10.7 percent of commuters on these relations use the high-speed train as a mode of transportation on their way to work.

Overall, the evidence suggests that reductions in travel time in the German railway network have not only led to an increase in commuter numbers of about 91,500 between 1999 and 2010, but have also raised the share of workers commuting to work by train. In particular, high-speed trains as one specific mode of transportation exhibit a substantial potential to support a modal switch towards the use of trains as a means of commuting. The estimates suggest that direct high-speed trains have reached an average market share among commuters between 10 and 12 percent, which is two to three times larger than the share of rail commuters on all other relations. In a report for the European Commission, Steer Davies Gleave (2006) shows that on medium distances, like on the connection between Frankfurt and Cologne, this market share can rise to more than 60 percent. This is consistent with our results by distance bands, which show that the effect of reductions in travel time on commuter numbers is largest in the interval between 200 and 550 kilometers. On these relations, high-speed trains have contributed to '*the dramatic recovery of the railway's market share over medium distances*' (Barrón, Campos, Gagnepain, Nash, Ulied, and Vickerman, 2009, p.16) due to their comparative advantage in terms of travel time compared to road or air transportation.

6 Conclusion

Drawing on the expansion of the high-speed rail network in Germany between 1999 and 2010 as a natural experiment, we have examined the effect that changes in travel times between regions have on the intensity of interregional commuting as well as on the locational choices that workers make regarding their region of residence and of work. Using the introduction of a direct HSR-connection between a pair of counties as an instrument for commuting time, we have shown that a reduction in travel time by one percent raises the number of commuters between regions by 0.25 percent. This rise in commuter numbers is mainly the result of workers changing jobs from larger to smaller cities, which have become more easily accessible as a result of faster train connections. These smaller cities benefit from investments in rail infrastructure by gaining access to a larger pool of qualified workers who, in turn, reap the gains from an improved commuting infrastructure by enjoying the amenities of urban life while at the same time finding a productive labor market match in the periphery. Our findings suggest that investments in rail infrastructure have led to a modal shift towards rail transportation, in particular between regions that were connected by a direct high-speed train.

A logical next step would be to conduct a more rigorous cost-benefit analysis of the large scale investments into high-speed rail infrastructure that were undertaken during the last decades in countries like Germany, Japan, France and, more recently, China. To date, the scarcity of data on the costs of building and maintaining the high-speed rail system renders a rigorous analysis out of reach, even when leaving aside the much more complex ecological dimension. The insights provided in this paper regarding the effect that investments in rail infrastructure have on the spatial organization of the labor market might, however, be taken as a first step in this direction.

References

Ahlfeldt, G. M., and A. Feddersen (2017): "From Periphery to Core: Measuring Agglomeration Effects using High-Speed Rail," *Journal of Economic Geography (forthcoming)*.

Ahlfeldt, G. M., S. J. Redding, D. M. Sturm, and N. Wolf (2015): "The Economics of Density: Evidence From the Berlin Wall," *Econometrica*, 83(6), 2127–2189.

Banerjee, A., E. Duflo, and N. Qian (2012): "On the Road: Access to Transportation Infrastructure and Economic Growth in China," NBER Working Paper 17897, National Bureau of Economic Research.

Barrón, I., J. Campos, P. Gagnepain, C. Nash, A. Ulied, and R. Vickerman (2009): "Economic Analysis of High Speed Rail in Europe," Informes 2009, Fundación BBVA.

Baum-Snow, N. (2007): "Did Highways Cause Suburbanization?," *The Quarterly Journal of Economics*, 122(2), 775–805.

Becker, U. J. (2016): "Das Kosten-Nutzen-Verhältnis in der Bundesverkehrswegeplanung: Wissenschaftlicher Anspruch und Auswirkungen in der Praxis," *Zeitschrift für Verkehrswissenschaften*, 87(1), 1–16.

Chandra, A., and E. Thompson (2000): "Does Public Infrastructure Affect Economic Activity?: Evidence from the Rural Interstate Highway System," *Regional Science and Urban Economics*, 30(4), 457–490.

Costa, D. L., and M. E. Kahn (2000): "Power Couples: Changes in the Locational Choice of the College Educated, 1940 - 1990," *The Quarterly Journal of Economics*, 115(4), 1287–1315.

Davis, D. R., and D. E. Weinstein (2002): "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 92(5), 1269–1289.

Deutsche Bahn (2014): "Bahnzahlen," *mobil*, 8, 61.

Donaldson, D. (2016): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review (forthcoming)*.

Duranton, G., P. M. Morrow, and M. A. Turner (2014): "Roads and Trade: Evidence from the US," *Review of Economic Studies*, 81(2), 681–724.

Duranton, G., and M. A. Turner (2012): "Urban Growth and Transportation," *The Review of Economic Studies*, 79(4), 1407–1440.

Faber, B. (2014): "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *The Review of Economic Studies*, 81(3), 1046–1070.

Federal Statistical Office (2016): “Erwerbstätige nach benutztem Verkehrsmittel für den Hinweg zur Arbeitsstätte, 1991-2012,” Data Set provided by the German Federal Statistical Office.

Garcia-López, M.-À., A. Holl, and E. Viladecans-Marsal (2015): “Suburbanization and Highways in Spain when the Romans and the Bourbons still Shape its Cities,” *Journal of Urban Economics*, 85, 52 – 67.

Gibbons, S., T. Lyytikäinen, H. Overman, and R. Sanchis-Guarner (2012): “New Road Infrastructure: The Effects on Firms,” SERC Discussion Papers 0117, Spatial Economics Research Centre, LSE.

Givoni, M. (2006): “Development and Impact of the Modern High-Speed Train: A Review,” *Transport Reviews*, 26(5), 593–611.

Glaeser, E. L., J. Kolko, and A. Saiz (2001): “Consumer City,” *Journal of Economic Geography*, 1(1), 27–50.

Haller, P., and D. F. Heuermann (2016): “Job Search and Hiring in Local Labor Markets: Spillovers in Regional Matching Functions,” *Regional Science and Urban Economics*, 60, 125 – 138.

Hantschel, A., and A. Schlange-Schöningen (2014): “Öffentliche Aufträge in Deutschland,” Merkblatt, Auftragsberatungszentrum Bayern e.V.

Heckman, J. J., H. Ichimura, and P. Todd (1998): “Matching as an Econometric Evaluation Estimator,” *Review of Economic Studies*, 65(2), 261–94.

Helpman, E., M. Melitz, and Y. Rubinstein (2008): “Estimating Trade Flows: Trading Partners and Trading Volumes,” *The Quarterly Journal of Economics*, 123(2), 441–487.

Henneke, M. (2011): “Morgens mit dem Zug zu VW nach Wolfsburg, abends zurück in die Hauptstadt: Viele Berliner sind Langstreckenpendler - und loben die Bahn AG,” *Berliner Zeitung*, Electronic Issue, February 14^{th} 2011.

Heuermann, D. F., and J. F. Schmieder (forthcoming): “The Effect of Infrastructure on Worker Mobility: Evidence from High-Speed Rail Expansion in Germany,” *Journal of Economic Geography*.

Kostka, G., and N. Anzinger (2015): “Large Infrastructure Projects in Germany - A Cross-Sectoral Analysis,” Working Paper 1 in the Project ‘Large Infrastructure Projects in Germany Between Ambition and Realities’, Hertie School of Governance.

Lalive, R., S. Luechinger, and A. Schmutzler (2013): “Does Supporting Passenger Railways Reduce Road Traffic Externalities?,” ECON - Working Papers 110, Department of Economics - University of Zurich.

Martín, J. C., and A. Reggiani (2007): "Recent Methodological Developments to Measure Spatial Interaction: Synthetic Accessibility Indices Applied to High-Speed Train Investments," *Transport Reviews*, 27(5), 551–571.

Matha, T., and L. Wintr (2009): "Commuting Flows across Bordering Regions: A Note," *Applied Economics Letters*, 16(7), 735–738.

McArthur, D. P., G. Kleppe, I. Thorsen, and J. Ubøe (2013): "The Impact of Monetary Costs on Commuting Flows," *Papers in Regional Science*, 92(1), 69–86.

Melo, P. C., D. J. Graham, and R. B. Noland (2011): "The Effect of Labour Market Spatial Structure on Commuting in England and Wales," *Journal of Economic Geography*, 12(3), 717–737.

Michaels, G. (2008): "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System," *The Review of Economics and Statistics*, 90(4), 683–701.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2015): "Commuting, Migration and Local Employment Elasticities," NBER Working Paper 21706, National Bureau of Economic Research.

Persyn, D., and W. Torfs (2016): "A Gravity Equation for Commuting with an Application to Estimating Regional Border Effects in Belgium," *Journal of Economic Geography*, 16(1), 155–175.

Pütz, T. (2015): "Verkehrsbild Deutschland," BBSR-Analysen KOMPAKT 15/2015, Bundesinstitut für Bau-, Stadt- und Raumforschung.

Qin, Y. (2014): "'No County Left Behind?' - The Distributional Impact of High-Speed Rail Upgrad in China," mimeo.

Redding, S. J., and D. M. Sturm (2008): "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, 98(5), 1766–1797.

Redding, S. J., and M. A. Turner (2015): "Transportation Costs and the Spatial Organization of Economic Activity," in *Handbook of Regional and Urban Economics*, ed. by G. Duranton, J. V. Henderson, and W. C. Strange, vol. 5, pp. 1339–1398. Elsevier.

Steer Davies Gleave (2006): "Air and Rail Competition and Complementarity," Report for the European Commission, DG Energy and Transport.

Vogel, F.-R. (2009): "Projektsteuerung öffentlicher Hochbauprojekte," Ph.D. thesis, University of Kassel.

Appendix

Table A.1: Sample Construction

Cities merged with surrounding county	
Amberg & Amberg-Sulzbach	Koblenz & Mayen-Koblenz
Ansbach (Stadt) & Ansbach (Land)	Landau (Pfalz) & Südliche Weinstraße
Aschaffenburg (St.) & Aschaffenburg (Land)	Landshut (Stadt) & Landshut (Land)
Augsburg (Stadt) & Augsburg (Land)	Ludwigshafen am Rhein & Rhein-Pfalz-Kreis
Bamberg (Stadt) & Bamberg (Land)	München (Stadt) & München (Land)
Bayreuth (Stadt) & Bayreuth (Land)	Osnabrück (Stadt) & Osnabrück (Land)
Coburg (Stadt) & Coburg (Land)	Passau (Stadt) & Passau (Land)
Darmstadt & Darmstadt-Dieburg	Pforzheim (Stadt) & Enzkreis
Erlangen & Erlangen-Höchstadt	Pirmasens (Stadt) & Südwestpfalz
Freiburg (Br.) & Breisgau-Hochschwarzw.	Regensburg (Stadt) & Regensburg (Land)
Heidelberg & Rhein-Neckar-Kreis	Rosenheim (Stadt) & Rosenheim (Land)
Heilbronn (Stadt) & Heilbronn (Land)	Schweinfurt (Stadt) & Schweinfurt (Land)
Hof (Stadt) & Hof (Land)	Straubing (Stadt) & Straubing-Bogen
Kaiserslautern (St.) & Kaiserslautern (Land)	Trier (Stadt) & Trier-Saarburg
Karlsruhe (Stadt) & Karlsruhe (Land)	Ulm (Stadt) & Alb-Donau-Kreis
Kassel (Stadt) & Kassel (Land)	Würzburg (Stadt) & Würzburg (Land)
Counties without train station in county capital (county capital in brackets)	
Aurich (Aurich)	Rhein-Hunsrück-Kreis (Simmern/Hunsrück)
Rheingau-Taunus-Kreis (Bad Schwalbach)	Tirschenreuth (Tirschenreuth)
Schwalm-Eder-Kreis (Homberg(Efze))	Landkreis Offenbach (Dietzenbach)
Saale-Holzland-Kreis (Eisenberg)	Werra-Meißner-Kreis (Eschwege)
Heinsberg (Heinsberg)	Donnersbergkreis (Kirchheimbolanden)
Lüchow-Dannenberg (Lüchow (Wendland))	Hohenlohekreis (Künzelsau)
Grafschaft Bentheim (Nordhorn)	Freyung-Grafenau (Freyung)
Saale-Orla-Kreis (Schleiz)	Sächsische Schweiz-Osterzgebirge (Pirna)
Vulkaneifel (Daun)	
Counties not contained in electronic schedules (county capital in brackets)	
Heidekreis (Bad Fallingbostel)	

Table A.2: Number of Observations per Year

Year	Number of Observations
1999	96,649
2000	99,267
2001	100,130
2002	98,514
2003	97,764
2004	98,352
2005	98,552
2006	99,610
2007	100,983
2008	101,315
2009	101,469
2010	102,128
Total	1,194,733

Table A.3: Cities with ICE-Stations in Rank Order (Population)

1991 - 1998	Inhabit.	Rank	1999 - 2010	Inhabit.	Rank
Berlin	3,461	1	Dresden	523	11
Hamburg	1,786	2	Leipzig	523	12
München	1,353	3	Aachen	259	29
Köln	1,007	4	Halle(Saale)	233	31
Frankfurt(Main)	680	5	Oberhausen	213	35
Stuttgart	606	6	Lübeck	210	36
Düsseldorf	589	7	Erfurt	205	37
Dortmund	580	8	Rostock	203	38
Essen	575	9	Saarbrücken	178	43
Bremen	547	10	Solingen	159	50
Hannover	523	13	Regensburg	135	56
Nürnberg	506	14	Ingolstadt	125	57
Duisburg	490	15	Wolfsburg	122	59
Bochum	375	16	Erlangen	105	72
Wuppertal	349	17	Jena	105	75
Bonn	325	18	Kaiserslautern	97	81
Bielefeld	323	19	Lüneburg	73	116
Mannheim	313	20	Bamberg	70	122
Karlsruhe	294	21	Aschaffenburg	67	125
Münster(Westf)	280	22	Weimar	65	130
Wiesbaden	276	23	Herford	64	136
Augsburg	266	24	Neustadt(Weinstr)	52	176
Braunschweig	249	27	Passau	49	185
Kiel	240	30	Lutherstadt Wittenberg	47	192
Magdeburg	232	32	Gotha	44	215
Freiburg(Breisgau)	224	34	Eisenach	42	245
Mainz	199	39	Stendal	40	251
Kassel	198	40	Siegburg	39	279
Hagen	189	41	Naumburg(Saale)	33	340
Hamm(Westf)	179	42	Bad Hersfeld	28	392
Potsdam	168	48	Köthen	27	436
Oldenburg	164	51	Saalfeld(Saale)	25	457
Heidelberg	147	53	Lichtenfels	20	664
Darmstadt	144	54	Montabaur	12	-
Würzburg	125	58		**117**	**119**
Ulm	122	60			
Offenbach(Main)	120	61			
Göttingen	120	63			
Koblenz	106	69			
Hildesheim	99	78			
Worms	80	100			
Neumünster	77	106			
Brandenburg	71	121			
Fulda	65	134			
Offenburg	57	150			
Baden-Baden	53	166			
	412	**28.5**			

The left panel contains all cities in Germany that have obtained a regular ICE-stop between 1991 and 1998; the right panel contains all cities where a regular ICE-stop was opened between 1999 and 2010; *Inhabit.* refers to population number in 2010; *Rank* denotes the rank that each city holds in terms of population among all 685 large and medium-sized cities in Germany, defined as those with more than 20,000 inhabitants; bold entries in the last row of each column contain mean population and median rank.

Table A.4: Propensity Score Matching

	Matching Regression, 1999	Comparison of Means		
Dependent Variable	ICE Introduction, 1999 - 2010	Treated	Untreated	t-statistic
$\log(W_{Orig})$	0.044 (0.069)	11.63	11.61	0.25
$\log(W_{Dest})$	0.398 (0.069)**	11.92	11.88	0.59
$\log(Population_{Orig})$	0.196 (0.062)**	12.56	12.55	0.76
$\log(Population_{Dest})$	-0.069 (0.059)	12.61	12.57	0.45
$\log(PopDensity_{Orig})$	0.193 (0.039)**	6.88	6.88	0.05
$\log(PopDensity_{Dest})$	0.086 (0.042)*	7.01	7.00	0.11
$\log(Distance)$	0.629 (0.075)**	5.34	5.28	0.97
$\log(Duration)$	-1.005 (0.078)**	4.89	4.83	1.08
$\log(Commuter\ Number)$	0.002 (0.023)	4.07	4.11	-0.33
Indirect ICE Connection	-0.065 (0.047)	0.36	0.39	-1.13
Pseudo R^2	0.31			
N	74,346			

* $p < 0.05$, ** $p < 0.01$; the second column contains the point estimates from a Probit regression of the probability that two counties are newly connected by a direct ICE-connection between 1999 to 2010; the model contains dummy variables for region type as well as state dummies (each separately for origin and destination county; point estimates not shown here); an *Indirect ICE Connection* is defined as a relation where an ICE-train is used on at least one leg of the trip but no direct HSR-connection exists; in the two rightmost columns the means of each variable are compared between treatment and control group.