

Heuermann, Daniel F.; Schmieder, Johannes F.

Conference Paper

Warping Space: High-Speed Rail and Returns to Scale in Local Labor Markets

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik
- Session: Local Labour Markets, No. A13-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Heuermann, Daniel F.; Schmieder, Johannes F. (2014) : Warping Space: High-Speed Rail and Returns to Scale in Local Labor Markets, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Local Labour Markets, No. A13-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100293>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Warping Space:

High-Speed Rail and Returns to Scale in Local Labor Markets*

Daniel F. Heuermann[†]
Regensburg University

Johannes F. Schmieder[‡]
Boston University, NBER, and IZA

February 2014

Abstract

Local returns to scale in the labor market have been notoriously difficult to disentangle from increasing returns in the product market and from the spatial sorting of workers and firms as a source for regional variation in productivity. In this paper we use the introduction of high-speed rail as a natural experiment in order to isolate the impact of labor market size on urban wages from product market and sorting effects. The key idea underlying our identification approach is that high-speed trains reduce commuting times between regions and thereby effectively increase the size of local labor markets without directly affecting product markets. The exact timing of the opening of high-speed rail connections can be regarded as exogenous, as a high-speed rail network is very expensive to build, requires a long planning phase and is mainly the result of political decisions. Furthermore, especially in the second wave of network expansion, several small towns were connected to the high-speed rail network simply because of their location between major metropolitan hubs and in this way got 'lucky' compared to neighboring towns. Drawing on a large and novel panel data set on the introduction of ICE-stations and on connection times between regions in Germany, as well as on a full sample of workers' employment histories, we examine the effect of high-speed trains on commuting behavior and wages. Using case studies, a pooled event study, and gravity equations with instrumental variables and propensity score matching we show that high-speed trains reduce traveling times by sixteen percent on average and significantly raise the number of commuters between local labor markets. We find that commuters incur wage gains of about three percent after the opening of an ICE-station, indicating that improved access to larger urban labor markets is associated with productivity gains for workers living in peripheral regions. In sum, our results suggest that between one third and half of overall agglomeration externalities are rooted in increasing returns to scale in local labor markets.

*We would like to thank David Card and Joachim Möller as well as seminar participants in Boston, Regensburg, Dortmund, Manchester, Hamburg, Trier and at the IAB for many helpful comments on this project. Peter Haller and Deborah Goldschmidt provided excellent research assistance. All errors are our own.

[†] Department of Economics, University of Regensburg, 93040 Regensburg, Germany (email: daniel.heuermann@wiwi.uni-regensburg.de; website: <http://www-wiwi.uni-regensburg.de/Personen/Daniel-F.-Heuermann.html.en>).

[‡] Department of Economics, Boston University, 270 Bay State Road, 02115 Boston, USA (email: johannes@bu.edu; website: <http://sites.google.com/site/johannesschmieder/>).

1 Introduction

It is a well established fact that workers earn higher wages in cities. For all countries examined so far empirical studies provide robust evidence for the existence of an urban wage premium of about five to ten percent (see, e.g., Glaeser/Maré 2001 and Yankow 2006 for the US, and Lehmer/Möller 2010 for Germany).¹ While higher nominal wages indicate a higher marginal product of labor in cities, the reasons for such higher productivity are less clear. Three classes of explanations are usually put forth. First, higher urban wages may result from a sorting of more productive workers and firms into cities. Secondly, the proximity of workers and firms may yield productivity advantages and, hence, increasing returns to scale in product and labor markets. Thirdly, geographic characteristics like natural resources or ports might lead to the founding of cities in economically beneficial locations and thereby be the underlying reason for higher urban productivity. A number of papers have tried to disentangle the relative importance of these mechanisms for the occurrence of the urban wage premium. The most recent contributions by Combes et al (2010) and Combes et al (2012) both find that agglomeration economies rather than geographic advantage or sorting are the driving forces behind higher urban productivity.

The prominent role of agglomeration economies as an explanation for higher urban wages raises the question whether increasing returns to scale mainly arise in the labor market or in the product market. With respect to product market effects, lower transportation costs, the availability of non-tradable specialized intermediary inputs and a positive selection of firms due to more intense competition might be the driving forces behind higher urban productivity. In contrast, agglomeration effects can also arise in the labor market as a result of a more efficient matching process between workers and firms in thick labor markets and because of knowledge spillovers.

Empirically, it has proven difficult to distinguish between these two channels since both are inherently correlated. Ideally, a natural experiment could be found where an exogenous change in the size of one of the markets is not accompanied by a change in the other one. The current literature has employed natural experiments in order to gain insight into the size of agglomeration effects but does usually not differentiate between product and labor market channels (see e.g., Davis/Weinstein 2002 and Greenstone/Hornbeck/Moretti 2010). A notable exception is the contribution by Redding/Sturm (2008) who draw on the case of German division and reunification as a natural experiment in order to examine the effects of changes in product market access on city growth. In line with predictions from New Economic Geography models they find that the spatial distribution of city sizes can be explained by differences in market access. While being instructive on the role that market access plays for population growth, their results need not necessarily arise from product market access as commuting linkages for workers may also have been cut off by the inner-German border.

In this paper we provide a novel approach to disentangling agglomeration economies by using the introduction and expansion of the high-speed rail network in Germany as a natural experiment in order to isolate local increasing returns to labor market size. The key idea behind this approach is that a reduction in commuting costs due to faster train connections effectively increases the size of the labor market without affecting the product market. The German high-speed train, the InterCityExpress (ICE), was introduced in 1991 and has since then been used only for passenger transport.² Hence, the ICE allows workers to commute over longer distances while leaving the relative distance between

¹Heurmann/Halfdanarson/Südekum 2010 provide an overview of the urban wage premium literature.

² According to information provided by the Deutsche Bahn, 76 million passengers have used the ICE in 2012.

product markets unaffected.³ The high-speed rail network in Germany was introduced in two major waves. During the 1990s, all major German cities were connected by high-speed trains. After the turn of the century, the network was in a second wave extended to also cover smaller and medium-sized cities. This second wave is of interest for our identification approach because the establishment of a high-speed rail stations was often the result of political decision rather than economic consideration. As a result, a number of 'lucky' cities were in this phase endowed with a high-speed rail station while other comparable cities were not. In these cases the exact location and the timing of the opening of an ICE-station can be regarded as exogenous, as they were over long periods of time subject to political haggles between national, regional and local authorities and influenced by petitions and public decisions. A prime example is Montabaur, a small town with slightly above 12,000 inhabitants, which was connected to the high-speed rail network in 2002 as part of the new high-speed connection between the two metropolitan regions Rhein-Ruhr and Rhein-Main. The opening of an ICE-station was not only the result of a political compromise after a decade-long discussion (see Ahlfeldt/Feddersen 2012), it also significantly affected the commuting decisions of workers. Figure I compares the number of workers commuting out of the local labor market that Montabaur is located in (*Westerwaldkreis*) to a synthetic control unit consisting of a linear combination of all other counties (Abadie/Gardeazabal 2003, Abadie/Diamond/Hainmueller 2011). After four years of identical development, commuter numbers start to diverge in 2003, i.e., the year following the opening of the ICE-station in Montabaur. This gap increases further over time and in 2010 the number of out-commuters from Montabaur exceeds those in the synthetic control unit by about 2,000.

In our identification approach we make use of this quasi-random distribution of shocks to labor market size which occurred as a result of the second wave of high-speed rail network development in order to disentangle local returns to scale in labor markets from returns to scale in product markets. We draw on a novel data set which contains information on the establishment of ICE-stations, as well as on connection frequency and commuting times between county pairs, and combine this information with individual employment data for all workers in Germany between 1975 and 2009. Since these data also contain information on a worker's place of residence and place of work we are able to observe commuting behavior between counties on the aggregate as well as on the individual level. Combining both data sets allows us to examine the impact that a sudden shock in labor market size has, first, on the commuting behavior and, secondly, on the productivity of workers.

We proceed in two steps. Focussing on county pairs as the unit of analysis we first use case studies, a pooled event study and gravity equations with instrumental variables and propensity score matching in order to examine whether the introduction of high-speed rail has reduced traveling times between regions in Germany and, in turn, has led to a larger number of commuters between local labor markets. In a second step, we examine the productivity consequences of changes in the relative size of local labor markets. We construct two indicators which relate local labor market size to railway accessibility of regions. Exploiting the expansion of the high-speed rail network as an instrument for regional labor market size we estimate augmented individual wage equations in order to shed light on the relative importance of labor market size for the urban wage premium.

We find that the introduction and expansion of the ICE-network has reduced commuting times on average by sixteen percent. Workers respond to this reduction in travel times by a higher commuting

³The idea to also use the ICE for freight trains has been considered but was discarded in an early stage of technology development.

intensity. We find a time-elasticity of commuter numbers of .4 in the short run and .6 in the long-run, indicating that the number of commuters rises between .4 and .6 percent if connection times between regions fall by one percent. The estimates on the size of local returns to labor market size show that wages of commuters rise by about three percent compared to non-commuters, indicating that about half of the overall size of agglomeration externalities can be attributed to increasing returns in the labor market.

In the next section we describe the data and outline the introduction and expansion of the high-speed rail network in Germany. In section four we first examine the effects that the introduction of high-speed rail has on connection times and, in turn, on the commuting behavior of workers; we then estimate the size of local returns to labor market size using the expansion of the high-speed rail network as an instrument. Section five summarizes the results and provides an outlook.

2 Data and Descriptives

2.1 Data

We draw on three different data sets. First, we make use of the electronic train schedules that have been published on CD ROMs annually from 1993 onwards by the Deutsche Bahn AG, the German railway company. These schedules contain information on all passenger train connections between all 5,400 train stations in Germany. From these CDs we have extracted connection times between all 390 county capitals in Germany for the years 1993 to 2012. For each year we have taken the first Monday in June that is not a holiday and have calculated trip duration, the number of train changes, and the types of trains for the fastest connection between all 76,245 pairs of cities in the time window between 6:00 am and 9:00 am, i.e., a typical commuting time. The resulting data set contains information on the size of the commutable labor market and can be used to calculate the number of jobs that can be reached within a certain time (e.g., 30 min, 60 min, 90 min) of train commute. In addition, we have extracted the year of establishment of an ICE-station within a county as well as information on the existence of a direct ICE-connection between each county pair. The frequency by which single stations or connections are served varies substantially. As we focus on commuter relations, throughout this paper we define ICE-stations and ICE-pair-connections as those which are served by ICE-trains at least every four hours.

Information on regional characteristics like county area, population density, unemployment, GDP, land prices, longitude and latitude of the county capital, and natural endowments are provided by the Federal Statistical Office (Statistisches Bundesamt). Throughout the paper we define a region as a city or county (*Kreise und Kreisfreie Städte*). These units, which are equal to NUTS-III regions and contain on average about 200.000 inhabitants, are not necessarily self-contained labor markets. Rather, one core city is often surrounded by a number of smaller and more rural counties and it crucially depends on the existence of road and rail infrastructure to which extent workers from peripheral counties are able to commute to core cities for work.

The information on trip duration, ICE-connections and regional characteristics are merged with administrative data provided by the Institute for Employment Research (IAB), which contains the employment histories of all workers in Germany between 1975 and 2009 as well as individual information on wages, education, age, sector, tenure, and experience. Wages, which are measured as inflation

adjusted gross daily wages, are top-coded at the threshold of maximum social security payments. We account for this fact by means of imputation in the analysis below. The information on place of residence and place of working contained in the data allows for identifying mobility patterns on individual as well as on aggregate level. From this full sample of workers we have generated an annual panel data set on full-time employed workers. Table I provides descriptive statistics on wages, commuting distance and labor market size for workers of different commuter status. The average wage of all workers is 81 Euros. With an average wage of 76 Euros non-commuters earn slightly less, while commuters receive 92 Euros on average. Workers commuting between cities that are connected by a direct ICE receive on average 102 Euros. Workers commute on average 28 kilometers to work; if place of work and place of residence are connected by a direct ICE, average commuting distance rises to 132 kilometers. As we can only measure distance between the train stations of the core cities within counties, these numbers are likely to be overestimated. They are, however, highly instructive in relative terms as they show that the commuting distance of workers commute between cities that are connected by a direct ICE is on average fifty percent larger than the average commuter distance (87 kilometers). With respect to labor market sizes the numbers show that workers systematically commute from smaller into larger counties. This difference is particularly pronounced for ICE commuters, where the county of work is on average 56 percent larger than the county of residence. Specifically, with a reduction in travel time commuting distances as well as wages of commuters rise; in addition, workers commute to relatively larger labor markets. Before turning to the estimation approach we briefly outline the introduction and expansion of the German railway network, which is of key importance in our identification approach.

2.2 High-Speed Rail in Germany

As in a number of developed countries like Japan, France, Spain and Belgium, transportation policies in Germany have undergone a re-focus from motorized individual transport back to railway transportation in the late 1970s and 1980s, a shift often referred to as the “renaissance of the train”. After two decades of publicly funded research into tracks and propulsion technology, the high-speed train network was opened in Germany in 1991 and has since then steadily been enlarged. Technically, high-speed trains are for most parts running on the same tracks as other passenger trains. During the last twenty years, stations and tracks were successively either adjusted to the requirements of the ICE or in a smaller number of cases newly built for the exclusive use of high-speed rail traffic. Maps I to IV provide an overview of the spatial distribution of the stations that are serviced by ICE-trains and the timing of their opening. The high-speed rail era started with one major track running from Hamburg via Mannheim and Frankfurt to Munich, servicing a total of eleven stations. In 1997, this north-south axis was complemented by an east-west track running from Berlin to Cologne and covering a number of large cities in the Rhein-Ruhr area. In addition, the branches to Bremen and Freiburg were added to the network, so that in 1998 a total of 34 ICE-stops were in operation. At this point, all major cities in Germany were connected to the ICE-network and the number of direct ICE-connections amounted to 314. The second wave of the introduction of high-speed rail started with the opening of the Berlin-Munich line in 2000, when a number of smaller cities in Eastern Germany and Bavaria, including Jena, Naumburg, Saalfeld and Lichtenfels, were connected to the high-speed network, leading to a total number of 60 ICE-stations being serviced in 2005. Since then, the network has mainly been

extended by adding branches to peripheral cities, e.g., to Lübeck and Rostock. Currently, 70 counties are connected by 260 high-speed trains on at least four hour intervals, in most cases every one or two hours. The number of direct ICE-connections during this second phase has risen by 544 to a total number of 859 in 2009.

With the introduction of high-speed rail, connection times have decreased substantially. Figure II summarizes the development of average connection times between all 76,245 county pairs and the share of fastest connections that encompass the use of an ICE-train on at least one leg of the trip. In 1993, average travel time between a random pair of cities amounted to five hours and fifteen minutes. Until 2010, this duration has decreased by about nine percent to four hours and 47 minutes. During the same period, the share of connections which encompass the use of an ICE-train has risen from 48 to 75 per cent. Traveling time and ICE-share exhibit an inverse relation, indicating that significant reductions in the average connection time occurred mostly at points in time when major ICE-lines were opened.⁴

Not all regions have, however, benefitted to the same extent and at the same time from high-speed rail. Map V contains the average travel time needed within in each county to get to any of the 389 other counties. In 1994, when the main north-south ICE-line was already established, the regions that were best connected were those located on the corridor Hamburg - Frankfurt/Nuremberg - Munich. During this phase large cities were connected to each other out of economic considerations. Peripheral regions, especially those in the eastern part of the country, remained at this point largely unaffected from the introduction of high-speed rail. The second phase of network expansion was, in contrast, characterized by the opening of ICE stations in small and medium-sized cities located between larger metropolises. In these cases the reasons for being connected to the network were mainly political ones, as the states (*Bundesländer*) situated en route were obliged to contribute to the building costs and in return demanded the establishment of stops on their territory. Map VI contains the changes in travel times between 1999 and 2010 in absolute values. It shows that first and foremost peripheral counties in the eastern part, in Bavaria, and around the Rhein-Ruhr-area, which were connected to high-speed rail by the east-west line and the Berlin-Munich line, saw reductions in average travel times, in some cases of more than one hour. Table II contains a list of all ICE-stations that were opened between 1999 and 2009 and the rank of each city with respect to the number of inhabitants. Most of these ICE stations are located in 'lucky' cities, which were connected to the high-speed rail network for quasi-random reason and which as a result saw substantial reductions in average travel times. Prime examples of small cities experiencing dramatic reduction in commuting times are Montabaur, Limburg, and Ingolstadt, which with the introduction of high-speed rail were suddenly connected to large labor markets in Cologne (Limburg), Frankfurt and Bonn (Montabaur) and Munich (Ingolstadt).⁵ None of the cities which were endowed with an ICE-station between 1999 and 2009 ranks among the largest top ten and only two are among the first twenty (Leipzig and Muenster). The median rank is 116

⁴ Prominent cases for such reductions in commuting times are the opening of new stations for the EXPO 2000 (World's Fair) in Hanover and the introduction of the ICE-connection between Berlin and Munich in 1999. Similarly, the nine-minute drop in average travel times occurring between 2002 and 2003 partly goes back to the new high-speed track between Frankfurt and Cologne, which also connects Montabaur and Limburg to the ICE-network.

⁵The cities Montabaur and Limburg are located in the counties *Westerwaldkreis* and *Limburg-Weilburg*, respectively. Throughout the analysis we take the counties rather than the cities within them as our units of observations, but refer to both as *Montabaur* and *Limburg* as these are not only the names of the core cities within the counties but also of the two ICE-stations.

and the average number of inhabitants amounts to 111,000. These cases that are of prime interest here, as for these cities and the counties surrounding them the opening of a high-speed railway station constitutes a large and exogenous shock to the labor market.

3 High-Speed Rail and Returns to Scale in Local Labor Markets

3.1 High-Speed Rail and Commuting Behavior

In this section we address the question whether and to what extent the introduction of high-speed trains has led to a reduction in commuting times and, in turn, to a rise in commuter numbers. We begin with three case studies where the introduction of a direct ICE-connection has reduced travel time to an extent that two formerly distinct labor markets have de facto become one. These cases are also the most clear-cut ones in terms of exogenous variation in accessibility as the affected counties got lucky by receiving an ICE-station for political reasons. In a pooled event study we then compare the development of commuter numbers between counties which have seen substantial reductions in commuting times to cases where commuting time has remained constant. These two analyses are of a descriptive nature and serve to provide an understanding of the relationship between travel times and commuting numbers and to test the suitability of a direct ICE-connection as an instrument for traveling times. We then examine the causal effect of travel times on the number of commuters in a gravity equation framework where we use the existence of a direct ICE-connection between two counties as an instrumental variable. We corroborate these results by means of propensity score matching, where we compare the development of commuting times and commuter numbers on 'treated' and 'untreated' county pairs which in 1999 had the same probability of being endowed with a direct ICE-connection in the decade after.

3.1.1 Case Studies and Pooled Event Study

The three cases we look at are close to a 'perfect' experimental design because the ICE-stations were built mainly for political reasons in peripheral regions and, as a result, connection times have dropped to an extent that workers have suddenly gained access to large urban labor markets within a realistic daily commuting time. One of the most definite cases in this respect is the connection between Montabaur and Siegburg (Bonn), where the ICE-stations and the tracks were newly built. The track was opened in July 2002 as one key part of the new railway connection between Frankfurt and Cologne, which is part of the trans-European transport network (TEN-T) policies of the European Union. The key idea of the project was to better connect the two largest German metropolitan areas, the Rhein-Ruhr area (in the state of North-Rhine Westphalia) with the Rhein-Main area (mostly in the state of Hesse). The track is used exclusively by passenger trains, i.e., travel times for freight trains have remained constant and, hence, product market access between both regions has remained unchanged. As the state Rhineland-Palatinate, through which the tracks also run, demanded the establishment of a stop on its territory in return for its financial contribution, the town of Montabaur with its 12,000 inhabitants suddenly saw itself in the lucky position to be connected to the high-speed rail network. As a result, the connection time from Montabaur to Siegburg (Bonn) fell by three hours to just eighteen minutes. As shown in Figure IV, the number of commuters has since then increased

steadily from about 300 to 450. The newly built tracks also connect Siegburg (Bonn) to the city of Mannheim, which is located in the Rhein-Main area. The connection time between these two cities has dropped from three hours to about seventy minutes as a result of the direct ICE-connection. In the years after, the number of commuters has gone up from below ten to about forty. Similarly, the cities of Nuremberg and Ingolstadt were connected by a newly established high-speed route in 2005. In this case, Ingolstadt got lucky by being endowed with an ICE-station because it was located en route of the high-speed track between Nuremberg and Munich. As a result, commuting time between Nuremberg and Ingolstadt has decreased from about ninety minutes to below thirty minutes. Since then, the number of commuters has doubled from about 90 to nearly 180.

All three cases have in common a drastic reduction in traveling times resulting from the introduction of a direct high-speed train connection. The rise in commuter numbers following the opening of the connection is stunning. While the absolute numbers may appear moderate, one needs to consider that given the costs of commuting, traveling such distances on a daily basis is only feasible for high-earning workers. As the rise in absolute numbers is compatible with the transport capacity of one or two ICE-trains and as in all three cases commuting time and costs are substantially higher by car than by train, it is very likely that most workers are using the ICE-train for commuting to work. By means of examples, a yearly ticket from Montabaur to Bonn costs about 3,000 Euro per year. At the same time, this is much more economical than doing the same trip by car, which takes about 50 minutes and costs on average about 8,000 Euro per year. As such, ICE-trains provide an efficient option mainly for highly productive and mobile workers who commute from small, peripheral cities to large urban labor markets.

In a next step we examine the effect of a direct ICE-connection on commuter numbers by means of a pooled event study. The objectives of this analysis are twofold. First, we address the question whether the opening of a direct ICE-connection is generally followed by a rise in commuter numbers. Second, for the introduction of a direct ICE-connection to be an exogenous instrument for commuting times it needs to be the case that the effect on commuters only arises through reduced connection times and is not the result of a general preference for the ICE as a mode of transportation. We address both issues by comparing county pairs where trip duration has fallen as a result of the introduction of a direct ICE-connection to control pairs, where traveling time has remained constant despite the introduction of an ICE. Both groups are defined such that travel time ranges between 90 and 240 minutes in 1999. This restriction ensures that the two labor markets are not within daily commuting distance at the beginning of the observation period while they are also not too distant to attain a commuting relation by means of a direct high-speed rail connection. The treatment group consists of pairs of counties where a direct ICE-connection has been introduced and where travel time has gone down by more than 50 minutes. We compare this group to a control group of county pairs where a direct ICE-connection was introduced but where connection time has decreased by less than five minutes within the period of observation. Given these definitions we obtain 76 county pairs in the treatment group and 76 pairs in the control group. Figure VI compares average travel time and commuter numbers within both groups. At the beginning of the period, the average connection time ranges at 160 minutes for both groups. In the following years it falls from 165 minutes in 1999 to slightly below ninety minutes in 2009 in the treatment group while remaining constant in the control group. The bulk of this reduction in the treatment group occurs between 2002 and 2003, probably as a result of the newly built tracks between the Rhein-Ruhr-area and the Rhein-Main area. With a time lag of two years the average

number of commuters in the treatment group rises from about 140 to 175 in 2009. In the control group, the average number of commuters remains largely unchanged. While it follows the same overall time trend in the second half of the period it does not show the same upward tendency as commuter numbers in the treatment group, indicating that the introduction of a high-speed rail connection leads to significant gains in commuter numbers only if it comes along with substantial reductions in travel times. In terms of identification, this insight is of key importance because it strongly supports the argument that the introduction of a direct ICE-connection is an exogenous instrument for commuting times in a regression of commuter numbers on connection times.

3.1.2 Gravity Model with Instrumental Variables

In this section we estimate the causal relation between commuting times and commuters using a gravity model of commuting. Gravity models, in general, are used to estimate the intensity of spatial interaction as function of market size and distance. While they have been extensively employed to model trade flows (see Anderson 2011 for an overview) and to a minor extent processes of international migration (see, e.g., Karemera et al. 2000), they have only recently been applied to the case of commuting (Mathä/Wintr 2007, McArthur et al 2013, Melo et al 2011, Persyn/Torfs 2012). While these studies all in general examine the effect of some type of spatial friction on commuting decisions, they differ in their object of analysis. Mathä/Wintr (2007) and Persyn/Torfs (2012) both focus on inter- and intra-national border effects using the cases of Luxemburg and The Netherland; McArthur et al (2013) examine the effect of monetary costs on commuting decisions. Closest to the approach pursued here is the contribution by Melo et al (2011), who investigate regional differences in the effect of distance on commuter numbers in England and Wales. Most relevant in the present context is their finding that the availability of railway infrastructure as measured by the number of railway stations per square kilometer is associated with larger commuting distances. We carry their analysis one step further by explicitly taking into account the specific connection time between two regions rather than looking only at the general availability of railway services. In order to do so we formulate the following gravity model

$$C_{ijt} = aR_{it}^{\alpha}W_{jt}^{\beta}D_{ijt}^{\gamma} \quad (1)$$

where the number of commuters C between regions i and j at time t is expressed as a function of labor market size at a worker's place of residence R_i and place of work W_j as well as of the trip duration D between the two locations i and j . Taking logarithms and including connection and time fixed effects c_{it} and t_t yields

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

The variable of interest here is the distance-decay gradient γ , which provides an estimate of the effect that the commuting time between a pair of cities has on the number of commuters between the two cities.⁶ An obvious problem with this specification is the endogeneity of traveling times, as the expansion of the rail infrastructure is likely to be influenced by changes in demand, i.e., by shifts in commuters. In order to address this issue we employ an instrumental variable approach where we use

⁶This parameter is sometimes also referred to as distance-deterrence parameter, see McArthur et al. (2010) and McArthur et al. (2013).

the introduction of a direct ICE-connection between a pair of cities as an instrument for travel time. For the instrument to be exogenous two requirements need to be fulfilled. First, the place and the timing of the opening of an ICE-station must not be influenced by changes in commuter numbers between a pair of cities. As argued above, for the second wave of high-speed rail introduction this assumption is plausible because virtually all cities that received an ICE-station during this phase were smaller cities located en route of an ICE-track that was either already established before 1999 and/or in order to connect two larger metropolises. In these cases the exact choice of an ICE-station was often the result of a political opportunism, an extensive public debate and topographical and geological considerations. Ahlfeldt/Feddersen (2012) provide a detailed description of the twists and turns that the political, legal and administrative process has gone through before the ICE-stations in Limburg and Montabaur were finally approved and built. In addition, given the complexity and the legal uncertainty of such large-scale projects the exact timing of the opening opening was over long periods hard to predict as in virtually all projects of this dimension legal cases on resident protection and on potential natural damaged were settled in court. As a result, whether and when the direct connection between two of these smaller cities was ultimately upscaled to high-speed rail traffic was largely independent from annual changes in commuter numbers. As a second requirement for instrument exogeneity, commuter numbers must only be indirectly influenced by a direct ICE-connection, namely its impact on traveling times. As the event study has shown commuter numbers do not rise with the mere existence of an ICE connection but only with a noticeable drop in connection times. Framed more simply, even given all the fascination that might emanate from modern transportation technologies, it would appear odd to believe that workers change their jobs across regions just to sit in an ICE with no time benefit. With these considerations in mind we employ the introduction of a direct ICE-connection as an instrument for traveling times and formulate the first stage as follows

$$\log(D_{ijt}) = b + \delta \log R_{it} + \theta \log W_{jt} + \lambda ICE_{ijt} + c_{ij} + t_t + \nu_{ijt} \quad (3)$$

Importantly, when estimating the model for the years 1999 and 2009 and including connection fixed effects, λ is identified only through ICE-connections that were established within this period and through the accompanying changes in connection times between pairs of cities. We estimate equations (2) and (3) for two different time intervals. First, we estimate the model for annual intervals, making use of the full panel structure of the data which covers all years between 1999 and 2009. In this case γ provides an estimate of the short-term adjustments, i.e., of the extent to which workers adjust their commuting behavior to changes in travel times within in the turn of one year. Not all adjustments are, however, likely to happen immediately after a change in the train schedules but rather occur over longer time horizons. We therefore also estimate both equations only for the years 1999 and 2010, i.e. over a time horizon of eleven years. In this case, the estimates can be interpreted as a long-run equilibrium where γ provides an estimate of the total effect that changes in commuting times have exerted on commuter numbers within the eleven years.

Across all regressions the problem of zero-entries, which with respect to trade matrices has received substantial attention in the literature (see, e.g., Helpman/Melitz/Rubinstein 2008), is of no relevance because all entries in the commuter matrix are strictly positive, i.e., between all pairs of cities there is at least of person commuting. On the downside this implies that the time elasticity of commuting can be estimated only with respect to the internal and not the external margin.

Table III displays the results of estimating equation (3) using annual data. In the first two columns

trip duration between two cities is expressed as a function of a direct ICE-connection. The results show that the trip duration between two cities that are connected by a direct ICE is on average 134 minutes, i.e., by a factor of four, longer than between two cities that are not connected by high-speed rail. This appears reasonable as by the very nature of the technology high-speed trains are predominantly used on long-distance connections. In Columns III and IV we estimate the full gravity model by including variables for market size and distance as well as year dummies. Here, the coefficient on ICE turns negative, indicating that for county pairs with a direct ICE connection travel time is by 51 minutes, or 60 percent, lower on average than for county pairs not connected by an ICE. As it is usually the case with gravity equations, the variation in the outcome variable that is explained by the model is very large. In the columns V and VI we include pair fixed effects are included. As a result, γ now only measures the time-effect of ICE-connections that were introduced between 1999 and 2009. The distance measure is dropped because it is time-invariant. The coefficient on ICE shows that a the introduction of a direct high-speed rail connection within the period of observation reduces travel times by about 12 minutes, which is equivalent to a reduction of ten percent. The fact that with pair-fixed effects this coefficient gets smaller shows that the time effect of an introduction of a direct ICE connection within in the period of observation is lower than that in the period prior to 1999, reflecting the fact that in this period the distances between the ICE-stations get substantially lower as the network gets denser and, hence, the time savings cannot be as large anymore as they used to be in the first wave.

Table IV shows the same specifications estimated only for the years 1999 and 2009. The results in columns I to IV remain virtually unchanged, indicating that the total time effect of an ICE-introduction is dominated by connections opened before 1999. In columns V and VI, where pair fixed effects are included and, hence, only the effect of the later wave of connections is estimated, the effect gets slightly bigger than in the annual panel. On connections where an ICE was introduced within the ten years of observation, trip duration fell by about 15 minutes, or 14 percent. Comparing these estimates to the results obtained from the annual panel indicates that the total time effect in the long-run equilibrium exceeds the short-term adjustment occurring within the first twelve months.

In sum, the results from the first stage show that the introduction of high-speed trains has made a difference for the relative proximity of local labor markets. Specifically, the opening of a direct ICE-connection has reduced the connection time between cities in Germany by between 10 and 14 percent. The fact that the coefficients are all highly significant provide first evidence that on the existence of a direct ICE connection between two cities is a relevant instrument for traveling times. At the same time, while being highly significant, the existence of a direct ICE connection explains only a minor part of the overall duration of a trip, which is not very suprising given the multitude of factors determining the length of a trip. However, the fact that the F-statistic in the first stage is larger than ten throughout all specifications shows that a direct ICE-connection is not a weak instrument for traveling times. With this result in mind we now use the introduction of a direct ICE connection as an instrument to estimate the time-elasticity of commuting decisions, i.e., we examine whether and to which extent workers respond to changes in travel times. We therefore estimate equation (2), using the specifications in columns V and VI in Tables III and IV as the first stage.

Table V contains the results obtained from the annual panel. In the first two columns we estimate a reduced form of the instrumental-variable model by regressing commuter numbers on the existence of a direct ICE connection. Column I contains the results from a univariate regression of commuter numbers on a direct ICE connection. It shows that the number of workers commuting between cities

connected by a direct ICE is about six times larger than between cities that are not connected by an ICE. When estimating the gravity model with fixed effects, the introduction of a new direct connection between 1999 and 2009 is associated with an increase in commuter numbers of 1.8 percent. Columns III to VI display the results from estimating the gravity model in equation (2) with and without instruments. The coefficient on trip duration varies substantially, depending on whether connection fixed effects are included or not. Comparing the regressions in columns III and IV, i.e., with and without fixed effects, shows that the coefficient on travel time picks up much of the the idiosyncratic characteristics of a connection. This is confirmed by the R^2 , which indicates that much of the variance in commuter numbers are explained by connection fixed effects. In columns V and VI we estimate the same equation with instrumental variables. Without fixed effects the results remains largely the same as in the OLS-regression. With fixed effects and instrumental variable, a reduction in commuting times is associated with an increase in commuter numbers of .3 percent. The difference in size can be rationalized by the insight from Maps V and VI, i.e., that reductions in connection times during the period of observation occurred mainly in peripheral regions with a lower number of commuters. Hence, when using instrumental variables we address the problem of reverse causality which stems from the fact that investments in rail infrastructure during the period of observation took mainly place in the periphery. As this is also the case with new ICE connections the effect of railroad infrastructure on commuter numbers is heavily underestimated even when using pair fixed effects. When addressing this selection effect by means of instrumental variables the results from the annual regressions indicate that a reduction in travel times by one percent is accompanied by a rise in commuter numbers by .3 percent.

Before turning to the estimates of the long-term model, the results obtained on the other variables in the gravity model warrant some comments. First, the coefficients on the size of the origin and destination region remain largely the same across all specifications. Roughly speaking, raising the number of workers in the origin and destination counties by one percent raises the number of commuters by .3 and .8 percent, respectively. This result is plausible in two respects. First, the relative magnitude of the coefficients pinpoints the relative importance of the demand side for commuting as compared to the supply side. Specifically, these numbers indicate that a rise in job in a region by one percent is associated with a probability of .8 percent that a worker from another regions starts to commute. The other way round, one percent more workers residing in a county only leads to a rise in out-commuters by .3 percent. Framed differently, the availability of jobs pulls workers into regions, while the supply of more workers in a region does only to a lesser extent push them into jobs in other regions. Secondly, the fact that both coefficients add up to about one reveals the general willingness to commute. If the size of the joint labor market, i.e. the sum of workers in origin and destination region, rises by one percent, then the number of commuters also rises by one percent. Hence, the share of workers that commute remains constant with the size of the joint labor market. The coefficient on distance varies between -.34 and -.48 in the OLS regressions. While this result is close to the size of .3 reported by Melo et al (2011), the substantial importance of the pair-fixed effects render any interpretation of this coefficients in simple OLS regressions rather useless.

As in the first stage, the effects get slightly larger in the long-run. As shown in Table VI, the effect of the introduction of a direct ICE-connection within the period of observation on commuter numbers in 2009 rises to .05, implying that on average five percent more workers commute between cities connected by a direct ICE compared to pairs that have were not endowed with a direct high-speed

rail connection. The coefficients on travel times vary in a similar way as before. The results from the instrumental variable regression with fixed effects indicates a time elasticity of commuter numbers of .05, i.e., a reduction of commuting time by one percent leads to a rise in commuter numbers by .5 percent.

As shown in the descriptives, only a small minority of connections is endowed with a direct ICE-connection. In the regression analysis we therefore have compared a small number of treated connections to a large number of untreated connections, which are likely to be very different in observable and unobservable characteristics. Assuming that observable and unobservable characteristics are correlated we therefore scrutinize our results by defining a sample of treated and untreated connections where we control for sorting on observables. We therefore define a sample of treated and untreated connections which are as comparable in terms of those observable characteristics which in 1999 are relevant for the later introduction of a direct ICE-connection. In order to obtain such a sample we conduct a propensity score matching with the introduction of a direct ICE-connection between 1999 and 2009 as the treatment variable. As predictors for this treatment we use population size and population density for both origin and destination counties, and connection time and commuter numbers in 1999 between the two counties. In addition, we use dummy variables for county type according to the classification by the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR) and employ fixed effects for county of destination and of origin. Table VII contains the results from the matching regression, where we choose a nearest neighbor approach. All variables turn out to be significant predictors for the later introduction of a direct ICE-connection. For a 517 of the 544 treated connections we are able to identify a nearest neighbor which has the same ex-ante probability of receiving an ICE-connection in the next ten years but failed to get one. The sample means and the accompanying t-statistics show that both sample are not significantly different any more in terms of observables. From these 1,035 county pairs we generate an annual panel for the years 1999 to 2009 and re-estimate equations (2) and (3) in order to examine whether our results hold when controlling for selection on observables. Columns VII and VIII in Tables III and IV show that the effects of infrastructure investments on travel time slightly increase in size as compared to the full sample. In the short run, the introduction of a direct ICE reduces connection time by 14 minutes, or eleven percent, and in the long-run by 19 minutes, or 16 percent. Similarly, the effects slightly increase in size when we re-estimate the effect of commuting times on commuter numbers, using the matched sample and instrumental variables in column VII in Tables V and VI. Here, a reduction on connection times raises commuter numbers by .36 percent in the short run and .6 percent in the long-run.

In sum, the results obtained in this part confirm the first prediction of our model, namely that a reduction in commuting costs between labor markets leads to a rise in the number of commuters. With respect to the case of rail infrastructure the regressions show that a reduction in connection times between regions by one percent leads to a rise in commuter numbers by .3 percent in the short run and by .6 percent in the long run. With this result we have made an important step towards identifying the returns to scale in local labor markets because we now can directly examine the productivity effects of changes in connection times.

3.2 Labor Market Size and Worker Productivity (*preliminary*)

In the prior section we have seen that workers take advantage of an enlargement of labor markets by an increased tendency towards commuting. In this section we address the productivity consequences of an enhanced labor markets by looking at the wages of commuters and non-commuters.

3.2.1 Case Study Evidence

As in the prior section we begin by examining one of the most clear-cut case in order to gain an understanding of the productivity consequences of an exogenous enlargement of the labor market. In order to identify the effect of a drop in connection times on commuter numbers we focus on the newly established station in Montabaur and examine whether gaining access to the high-speed rail network in 2002 is followed by higher local wages for commuters and non-commuters. We construct a sample of all workers who live in the county in 2002 and examine whether wages in Montabaur adjust to the new labor market situation and how these productivity gains are distributed between commuters and non-commuters. Constructing the sample like this avoids selection effects of workers who move to Montabaur shortly before the ICE-station is opened in order to commute to high-paying jobs. We therefore define a cohort of individuals, defined as all individuals who live in a county over the whole period of investigation, i.e., between 1999 and 2009. We then follow these individuals over time in order to examine whether workers incur wage gains after having access to a larger labor market. In a first step of this analysis we find that the number of workers commuting out of Montabaur to one of the eleven cities than Montabaur is connected to by a direct ICE rises by about fifty percent from 300 to 450 after between 2003, i.e., after the ICE-connection has been introduced. The average wage of workers commuting out of Montabaur increases sharply in 2003, namely by six percent from 132 to above 140 Euros. A regression analysis with individual fixed effects shows that about half of this rise is explained by the introduction of a direct ICE-connection to large labor markets. The wages of workers commuting to these eleven cities show a sudden positive level effect of about three percent, while the other three percent are explained by an overall positive time trend arising for all workers (the results currently are in the process of anonymity screening by the Research Data Centre and therefore cannot be provided in a table). These findings provide first evidence on the monetary returns of being connected to a larger labor market. Comparing this wage jump to the results by Lehmer/Möller (2010), who find for a genuine urban wage premium of about eight percent, provides tentative evidence that between one third and half of the urban wage premium is explained by increasing returns in the labor market.

3.2.2 Wage Regressions with Instrumental Variables

We corroborate the evidence from the case study by examining whether access to larger labor markets generally leads to higher wages of commuters and non-commuters. Theoretically, with an increased accessibility of potential jobs in adjacent labor markets we expect workers to be more likely to find more productive matches. To gain insight into matching quality in enlarged labor markets we estimate whether changes in the size of the commutable labor market are associated with significant wage gains. We therefore estimate Mincerian wage equations augmented by indicators for labor market size.

$$\log(w_{kt}) = c + X_{kt}\eta + Y_{ijt}\rho + \tau\log L_{jt} + \phi\log Dist_{ijt} + c_k + t_t + v_{kt} \quad (4)$$

Specifically, we estimate the wage of individual k at time t as a function of the characteristics of a workers X , of the pair of regions that a worker is living and working in Y , of the size of the labor market L , as well as as of the commuting distance between the place of residence and the place of working $Dist$. We include worker and time fixed effects throughout all regressions in order to control for unobserved worker heterogeneity as well as for common time-variant influences. The variable of interest here is L , which measures the size of the labor market a person is working in. As such, equation (4) has provided the standard identification approach in virtually all studies on the urban wage premium. So far, L has been measured by the number of workers or inhabitants within some administrative unit which is assumed to approximate a local labor market. The focus on single geographic units has been ubiquitous mainly because data are usually available for single administrative units and because a measurable source of exogenous variation in the interconnectedness of regions is hard to find. Despite its prominence, the difficulties with this standard approach are, first, that especially with an increased occurrence of commuting local labor markets never end at local borders and, second, that changes in labor market size are closely correlated with changes in the size of product markets.

We adress these shortcomings by proposing a novel definition of local labor markets which takes into account the interconnectedness of local labor markets and allows for disentangling changes in labor market size from changes in product market size. In order to do so we define two different indicators which both augment the size of the size of local labor markets a person is working in W_{jt} by a linear combination of the workers in all other counties weighted by travel time.

$$L_{jt}^1 = W_{jt} + \sum_{m=1}^{389} W_{mt}\Phi \quad \left. \begin{array}{l} \Phi = 0 \text{ if } D_{jmt} \geq 60 \\ \Phi = 1 \text{ if } D_{jmt} < 60 \end{array} \right\} \quad (5)$$

The measure L^1 expresses the size of a local labor market as the sum of jobs that can be reached within a daily commute, i.e., within sixty minutes of travel time from a worker's place of work. Note that in this definition labor market size does not depend on the Euclidian distance between two places, but on the travel time seperating them. In this definition the weighting is done in a binary way, i.e., all jobs within a daily commuting distance are counted equally, while those outside are ignored. This measure explicitly defines a local labor market as a central place for daily commuters. Connections between counties that are too distant from each other to allow for daily commuting are systematically ignored. In order to explicitly take into account long-distance commuting we define a second indicator L^2 which defines a the size of a local labor market as the sum of jobs all jobs in Germany weighted by travel time.

$$L_{jt}^2 = W_{jt} + \sum_{m=1}^{389} \frac{W_{mt}}{D_{jmt}} \quad (6)$$

The key difference between these two indicators is the relative emphasis they put on daily and long-distance commuting. The summary statistics for both indicators, which are contained in Table VIII, serve to illustrate this point. In order to control for reverse causality we again employ an intrumental variable approach. As in the prior section, we make use of the second wave of high-speed rail introduction, where mainly peripheral cities gained access to larger urban labor markets by means of direct ICE connections. Here, we use the number of direct ICE-connections as an intrument for L^1

and L^2 . The argument for instrument validity closely follows the discussion above. With respect to instrument exogeneity, it can plausibly be argued that the introduction of a direct connections between counties during the second wave of network expansion is largely the result of both political factors and complex processes of urban and regional planning. In addition, as shown above, a direct ICE connection influences the number of commuters only through time effects. The instrument can be shown to be relevant in the first stage regression where we regress the two indicators for regional labor market size on the same regional variables that are used in the second stage and on the number of direct ICE connections:

$$\log(L_{jt}^{1,2}) = d + Y_{jt}\varphi + \lambda ICE_{jt} + c_j + t_t + \mu_{jt} \quad (7)$$

We continue by estimating equation (4) with and without instrumental variables. In order to disentangle the effect that labor market size has for commuters and non-commuters we estimate the same specifications for all workers, for commuters, and for non-commuters. As the coefficients on individual characteristics are largely constant across all specifications and in line with findings from a large literature we refrain from commenting on them here. For the full sample of workers we find in a regression with fixed effects and instrumental variable that wages rise by between .18 and .26 percent with a increase in the size of local labor markets by 1 percent, depending on the indicator used. This in turn implies that an increase in the size of a local labor market by one standard deviation is associated with a rise in wages by between two and three percent (results currently are in the process of anonymity screening by the Research Data Centre and cannot yet be provided in a table). Given the finding of an overall wage premium of about eight percent, this implies that about one third of the total urban wage premium can be attributed to increasing returns to scale in local labor markets.

4 Summary and Outlook

The aim of this paper was to shed light on the extent to which agglomeration externalities are driven by increasing returns in the labor market. Addressing this question we have used the German case of the introduction and extension of a high-speed rail network as a natural experiment in order to disentangle increasing returns arising in the labor market from those arising in the product market. Focusing mainly on the introduction of direct ICE-connection during the second wave of infrastructure development, which occurred between 1999 and 2009 and has mainly affected peripheral cities, we have in a first step provided evidence that access to the high-speed rail network has reduced connection times between counties and has, in some cases dramatically, increased the accessibility of urban labor markets for peripheral counties. Using case studies, a pooled event study and combining gravity equations with an instrumental variable approach and propensity score matching we have shown that these reductions in average connection times are accompanied by sizable increases in the number of commuters. Specifically, we find that the introduction of a direct ICE-connection is followed by a rise in commuter numbers of about 1.8 percent. Using the introduction of a direct ICE connection as an instrument for commuting times provides evidence for a time elasticity of commuting of about .4 in the short and .6 in the long run, i.e., commuter numbers rise by between .4 and .6 percent with a reduction in travel time of one percent.

Turning to the productivity consequences of larger labor market size we have introduced two a novel

indicator for labor market size, which in contrast to standard approaches explicitly takes into account the potential of commuting between administrative units. Estimating augmented wage equations and employing the expansion of the high-speed rail network as an instrument for labor market size we provide evidence that for these commuters wages rise by about six percent, out of which half is explained by access to larger labor market. Taken together, these findings can be taken as evidence that about one third of the higher urban productivity is rooted in increasing returns to scale in local labor markets.

A logical next step would be to get somewhat closer to an idea of the cost-benefit ratio of the large scale investments into high-speed rail infrastructure that have been undertaken during the last two decades. The magnitude of the past investments as well as the question which type of mobility to support politically given an increasing number of commuters make this type of analysis a relevant endeavor. Even leaving aside the much more complex ecological dimension, the scarcity of data on costs of building and maintaining the high-speed rail system render this type of question not easy to answer. The first insights into the efficiency gains arising in the labor market provided in this paper might, however, be taken as a first step into that direction.

References

- [1] Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synth: Stata module to implement synthetic control methods for comparative case studies. Statistical Software Components, Boston College Department of Economics, October 2011.
- [2] Alberto Abadie and Javier Gardeazabal. The economic costs of conflict: A case study of the basque country. *American Economic Review*, 93(1):113–132, March 2003.
- [3] Gabriel M. Ahlfeldt and Arne Feddersen. From periphery to core: economic adjustments to high speed rail. Working Papers 2010/38, Institut d’Economia de Barcelona (IEB), 2010.
- [4] James E. Anderson. The gravity model. *Annual Review of Economics*, 3(1):133–160, 09 2011.
- [5] Gilles Duranton Laurent Gobillon Combes, Pierre-Philippe and Sébastien Roux. *Agglomeration Economics*, chapter Estimating Agglomeration Effects with History, Geology, and Worker Fixed-Effects, pages 15 – 65. Chicago University Press, 2010.
- [6] Pierre-Philippe Combes, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux. The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, 80(6):2543–2594, November 2012.
- [7] Donald R. Davis and David E. Weinstein. Bones, bombs, and break points: The geography of economic activity. *American Economic Review*, 92(5):1269–1289, December 2002.
- [8] Moshe Givoni. Development and impact of the modern high-speed train: A review. *Transport Reviews*, 26(5):593–611, 2006.
- [9] Edward L Glaeser and David C Mare. Cities and skills. *Journal of Labor Economics*, 19(2):316–42, April 2001.
- [10] Michael Greenstone, Richard Hornbeck, and Enrico Moretti. Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 118(3):536–598, June 2010.
- [11] James J Heckman, Hidehiko Ichimura, and Petra Todd. Matching as an econometric evaluation estimator. *Review of Economic Studies*, 65(2):261–94, April 1998.
- [12] Elhanan Helpman, Marc Melitz, and Yona Rubinstein. Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics*, 123(2):441–487, 2008.
- [13] Daniel Heuermann, Benedikt Halfdanarson, and Jens Suedekum. Human capital externalities and the urban wage premium: Two literatures and their interrelations. *Urban Studies*, 47(4):749–767, April 2010.
- [14] David Karemera, Victor Iwuagwu Oguledo, and Bobby Davis. A gravity model analysis of international migration to north america. *Applied Economics*, 32(13):1745–1755, 2000.
- [15] Florian Lehmer and Joachim Möller. Interrelations between the urban wage premium and firm-size wage differentials: a microdata cohort analysis for germany. *The Annals of Regional Science*, 45(1):31–53, August 2010.
- [16] Juan Carlos Martín and Aura Reggiani. Recent methodological developments to measure spatial interaction: Synthetic accessibility indices applied to high-speed train investments. *Transport Reviews*, 27(5):551–571, 2007.
- [17] Thomas Matha and Ladislav Wintr. Commuting flows across bordering regions: a note. *Applied Economics Letters*, 16(7):735–738, 2009.

- [18] David Philip McArthur, Gisle Kleppe, Inge Thorsen, and Jan Ubøe. The spatial transferability of parameters in a gravity model of commuting flows. Discussion Papers 2010/3, Department of Business and Management Science, Norwegian School of Economics, May 2010.
- [19] David Philip McArthur, Gisle Kleppe, Inge Thorsen, and Jan UbÅže. The impact of monetary costs on commuting flows*. *Papers in Regional Science*, 92(1):69–86, 2013.
- [20] Patricia C. Melo, Daniel J. Graham, and Robert B. Noland. The effect of labour market spatial structure on commuting in england and wales. *Journal of Economic Geography*, 2011.
- [21] Damiaan Persyn and Wouter Torfs. A gravity equation for commuting. Vives discussion paper series 33, Katholieke Universiteit Leuven, Faculteit Economie en Bedrijfswetenschappen, Vives, 2012.
- [22] Stephen J. Redding and Daniel M. Sturm. The costs of remoteness: Evidence from german division and reunification. *American Economic Review*, 98(5):1766–97, December 2008.
- [23] Christopher Taber and Hidehiko Ichimura. Propensity-score matching with instrumental variables. *American Economic Review*, 91(2):119–124, May 2001.
- [24] Jeffrey J. Yankow. Why do cities pay more? an empirical examination of some competing theories of the urban wage premium. *Journal of Urban Economics*, 60(2):139–161, September 2006.
- [25] Frankfurter Allgemeine Zeitung. Was 300 stundenkilometer kosten, 25. Juli 2012.

Appendix

Figure I: Out-Commuters from Montabaur

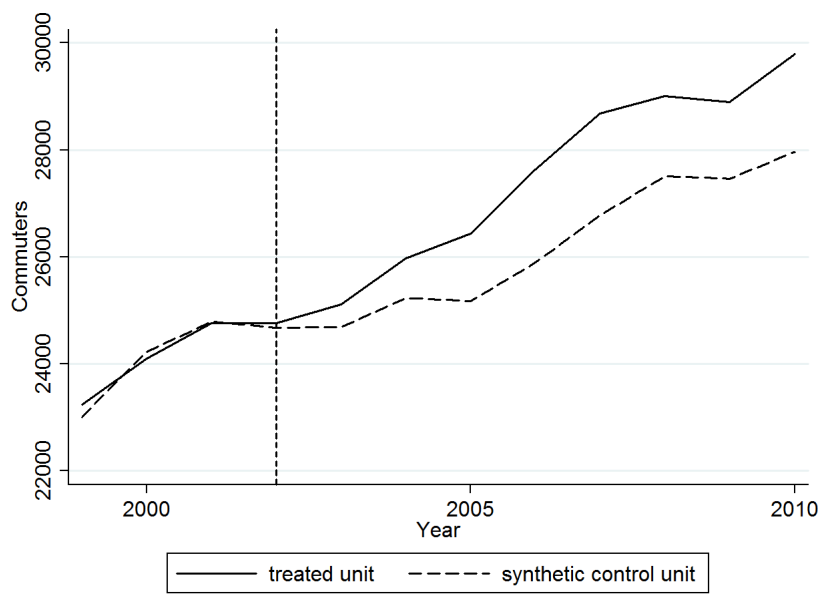


Table I: Descriptive Statistics of Worker Groups According to Commuter Status

	All Workers		Non-Commuters		Commuters		ICE Commuters	
Total Number (2009, in thousands)	22,918		15,598		7,320		337	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Wage	81.4	39.9	76.3	36.7	92.2	41.6	101.6	44.7
Distance	27.9	84.2	-	-	87.6	130.2	131.8	173.8
Size Origin	151,155	214,472	173,144	236,234	104,292	147,879	213,581	222,866
Size Destination	227,353	263,884	173,144	236,234	227,353	263,884	332,222	267,493

Maps I - IV: Evolution of the High-Speed Rail Network in Germany

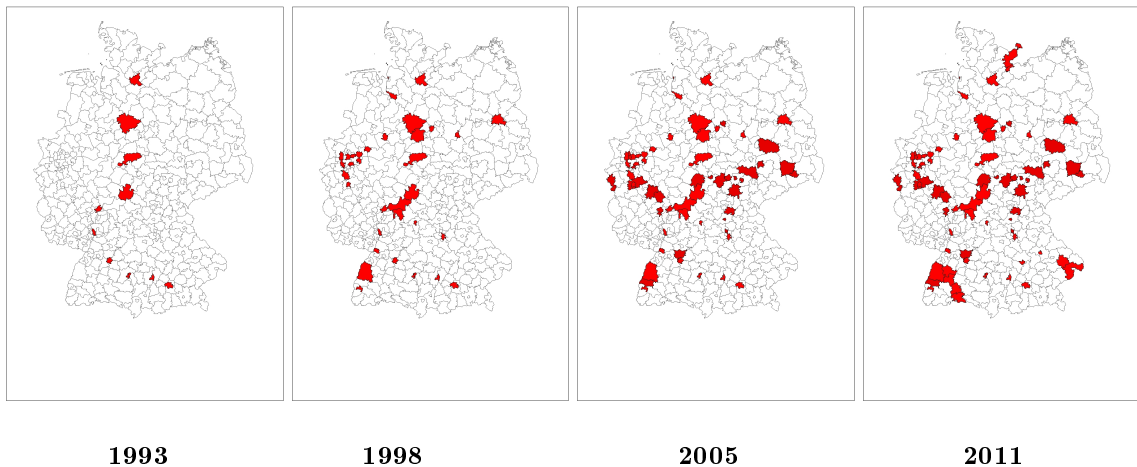
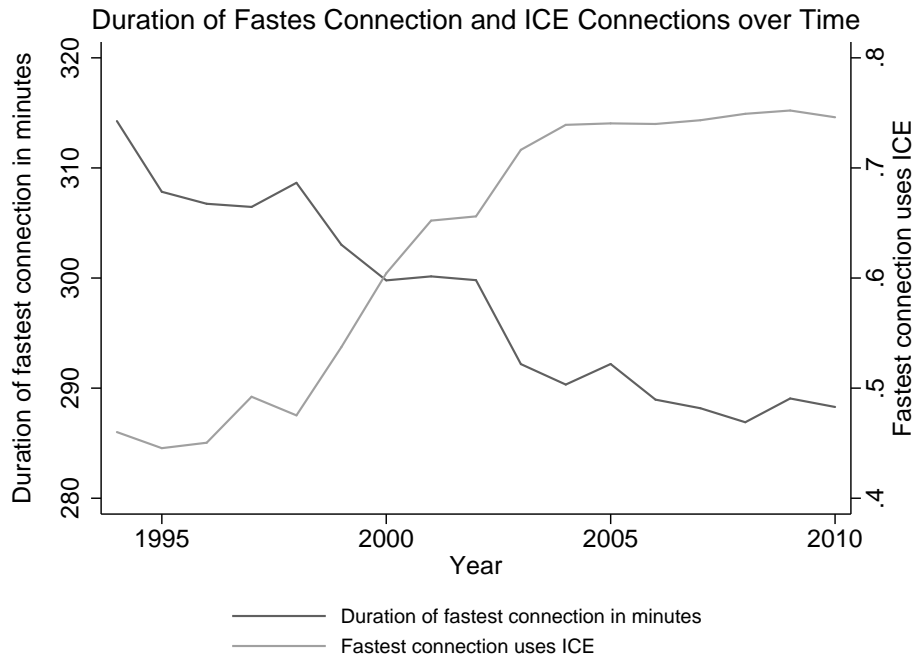


Figure II: Average Connection Times and Share of ICE-Trains



Maps V and VI: Regional Accessibility by Railroad

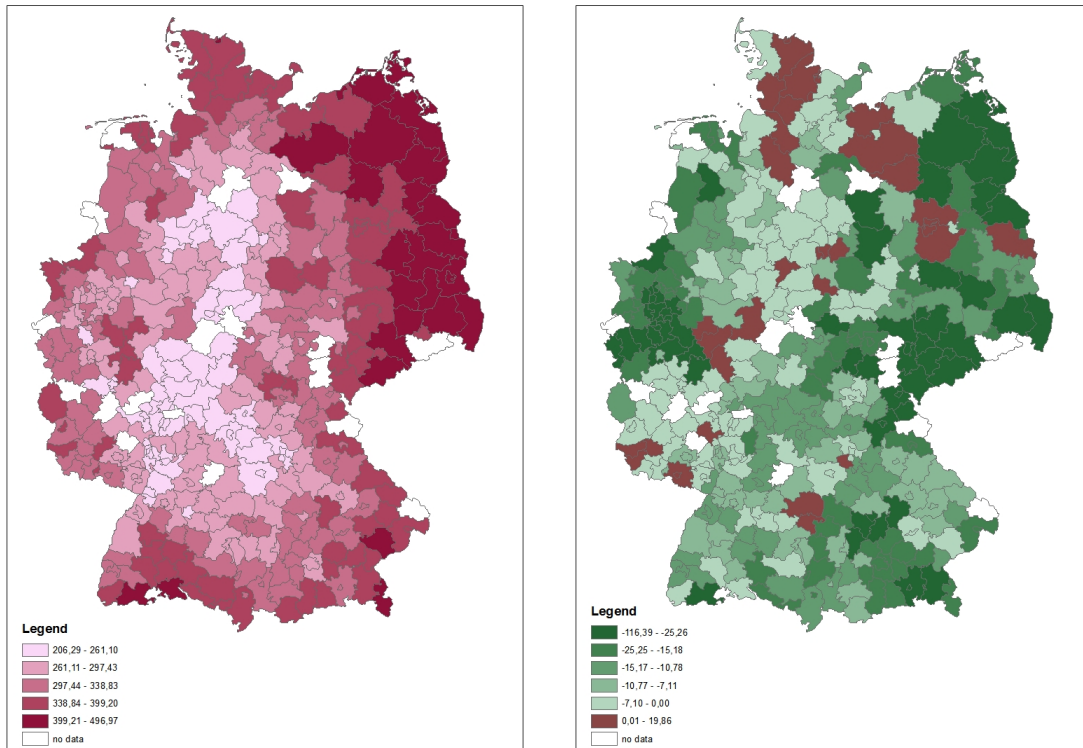


Table II: ICE-Stations Opened between 1999 and 2009 in Rank Order

Cities where an ICE Station was opened between 1999 and 2009	No of Inhabitants 2012 (in Thousands)	Rank
Leipzig Hbf	520	12
Muenster(Westf)Hbf	297	20
Aachen Hbf	240	29
Halle(Saale)Hbf	230	31
Oberhausen Hbf	210	35
Luebeck Hbf	211	36
Erfurt Hbf	203	37
Rostock Hbf	203	38
Saarbruecken Hbf	177	43
Osnabrueck Hbf	154	46
Solingen Hbf	155	50
Regensburg Hbf	138	56
Koblenz Hbf	108	71
Erlangen	105	72
Jena	107	75
Kaiserslautern Hbf	97	81
Guetersloh Hbf	95	84
Minden(Westf)	80	101
Lueneburg	70	116
Bamberg	71	122
Weimar	63	130
Herford	65	136
Neustadt(Weinstr)Hbf	52	176
Passau Hbf	49	185
Lutherstadt Wittenberg	47	192
Gotha	44	215
Homburg(Saar)Hbf	41	236
Eisenach	42	245
Siegburg/Bonn	39	279
Uelzen	33	339
Naumburg(Saale)Hbf	33	340
Limburg (Lahn)	33	349
Bad Hersfeld	28	392
Koethen	27	436
Saalfeld(Saale)	25	457
Salzwedel	25	517
Lichtenfels	20	664
Montabaur	12	-

Rank among all 685 large and medium-sized cities defined as those with more than 20,000 inhabitants

Data provided by the Federal Statistical Office

Figures IV - VI: Connection Times and Commuter Numbers (Case Studies)

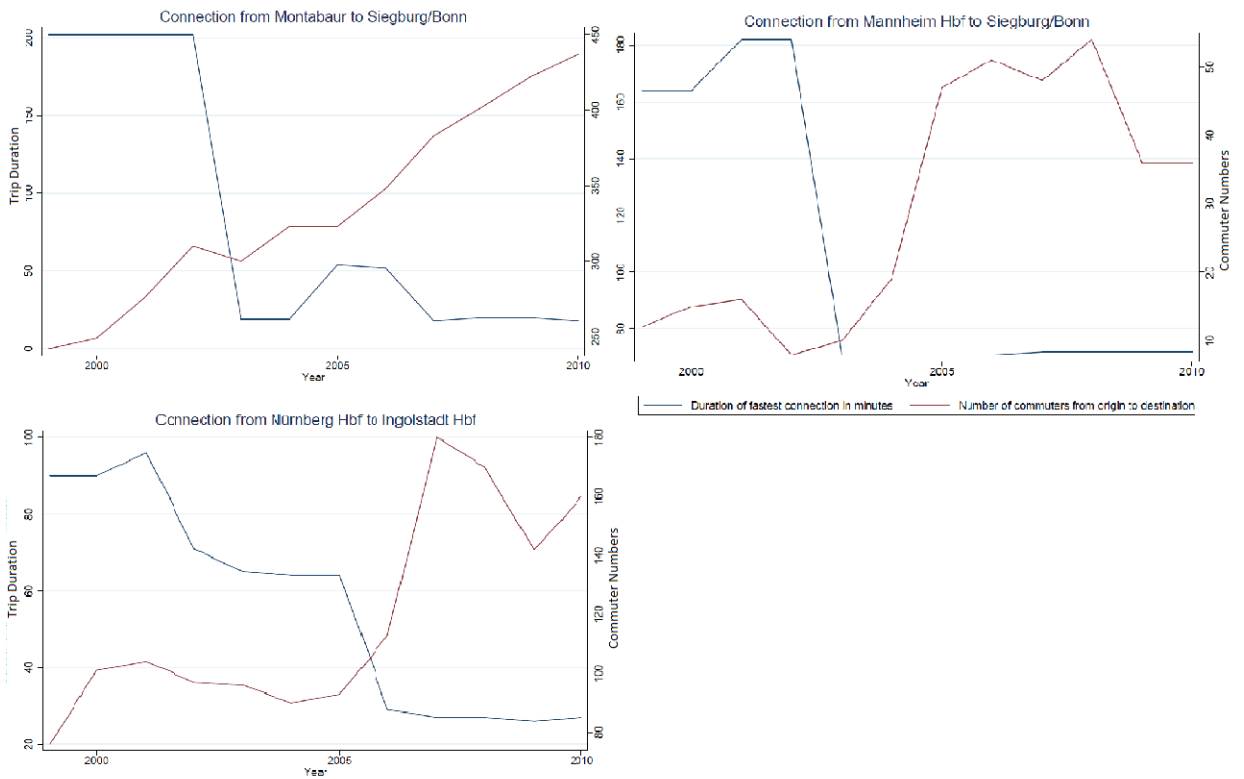


Figure VII: Connection Times and Commuter Numbers (Pooled Event Study)

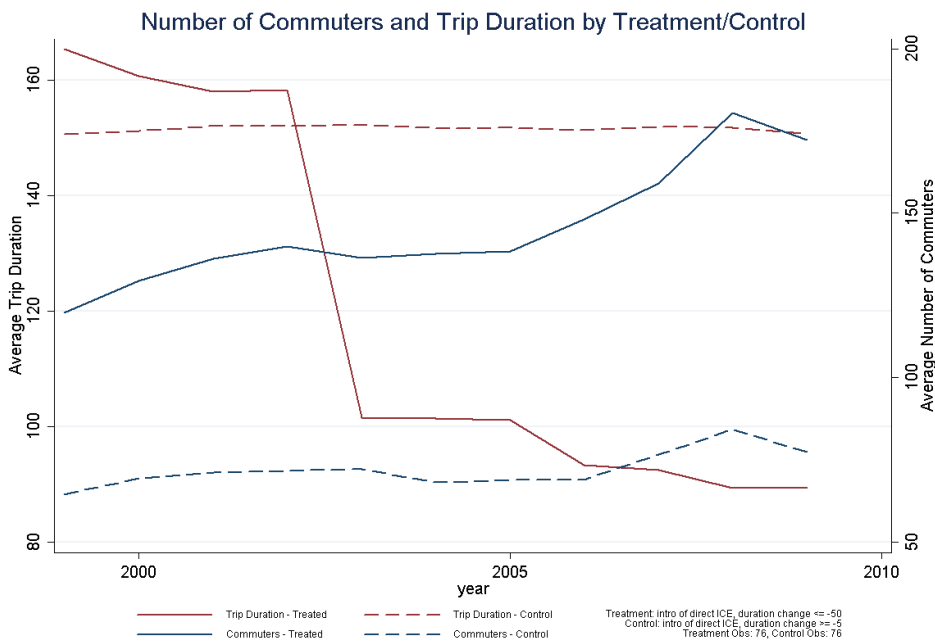


Table III: Direct ICE-Connection and Travel Times (Short-Run Adjustments)

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Dependent Variable:	Dur	log(Dur)	Dur	log(Dur)	Dur	log(Dur)	Dur	log(Dur)
<i>Direct ICE</i>	134.310 (3.065)**	4.615 (0.059)**	-51.256 (0.744)**	-0.601 (0.003)**	-12.073 (0.305)**	-0.093 (0.001)**	-14.114 (0.544)**	-0.113 (0.004)**
$\log(\text{Pop Orig})$			-21.720 (0.083)**	0.063 (0.000)**	35.790 (0.801)**	0.106 (0.004)**	10.419 (6.409)	-0.018 (0.047)
$\log(\text{Pop Dest})$			-23.272 (0.083)**	0.026 (0.000)**	8.601 (0.721)**	0.061 (0.003)**	-7.280 (0.532)	0.089 (0.038)*
$\log(\text{Distance})$			136.032 (0.111)**	0.773 (0.001)**				
R^2	0.00	0.01	0.94	1.00	0.11	0.08	0.00	0.11
N	683,369	683,369	683,230	683,230	683,230	683,230	11,053	11,053
Method	OLS	OLS	OLS	OLS	FE	FE	PSM, FE	PSM, FE

* $p < 0.05$; ** $p < 0.01$

Specifications in Columns 3 to 8 contain year-fixed effects.

Dependent variable is either Trip Duration (Dur) or the log thereof ($\log(\text{Dur})$).

Table IV: Direct ICE-Connection and Travel Times (Long-Run Equilibrium)

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Dependent Variable:	Dur	log(Dur)	Dur	log(Dur)	Dur	log(Dur)	Dur	log(Dur)
<i>Direct ICE</i>	133.452 (8.052)**	4.612 (0.155)**	-51.339 (1.948)**	-0.634 (0.009)**	-15.364 (1.107)**	-0.144 (0.005)**	-18.746 (1.999)**	-0.159 (0.015)**
$\log(\text{Pop Orig})$			-21.821 (0.195)**	0.065 (0.001)**	53.290 (1.784)**	0.169 (0.008)**	31.786 (13.623)*	0.153 (0.099)
$\log(\text{Pop Dest})$			-23.248 (0.194)**	0.030 (0.001)**	5.929 (1.655)**	0.074 (0.008)**	2.253 (10.940)**	0.184 (0.080)*
$\log(\text{Distance})$			137.634 (0.263)**	0.777 (0.001)**				
Year 2009			-11.937 (0.382)**	-0.029 (0.002)**	-14.048 (0.157)**	-0.053 (0.001)**	-11.407 (1.489)**	-0.054 (0.011)**
R^2	0.00	0.01	0.94	1.00	0.18	0.13	0.03	0.10
N	123,715	123,715	123,680	123,680	123,680	123,680	2,068	2,068
Method	OLS	OLS	OLS	OLS	FE	FE	PSM, FE	PSM, FE

* $p < 0.05$; ** $p < 0.01$ Dependent variable is either Trip Duration (Dur) or the log thereof ($\log(\text{Dur})$).

Table V: Travel Times and Commuter Numbers (Short-Run Adjustments)

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
Dependent Variable:	$\log(\text{Commuter Numbers})$						
<i>Direct ICE</i>	6.348 (0.038)**	0.018 (0.002)**					
$\log(\text{Duration})$			-1.110 (0.004)**	-0.019 (0.002)**	-1.264 (0.022)**	-0.317 (0.028)**	-0.368 (0.045)**
$\log(\text{Pop Orig})$		0.219 (0.010)**	0.330 (0.001)**	0.217 (0.010)**	0.420 (0.002)**	0.255 (0.011)**	0.398 (0.052)**
$\log(\text{Pop Dest})$		0.978 (0.008)**	0.547 (0.001)**	0.982 (0.008)**	0.612 (0.002)**	0.978 (0.009)**	0.915 (0.042)**
$\log(\text{Distance})$			-0.480 (0.003)**		-0.341 (0.017)**		
R^2	0.04	0.07	0.74	0.07	0.57	0.04	0.15
N	683,963	683,825	683,230	683,230	683,230	670,444	11,053
Method	OLS	FE	OLS	FE	IV	FE, IV	PSM, IV, FE

* $p < 0.05$; ** $p < 0.01$

Specifications in Columns 2 to 7 contain year-fixed effects

Table VI: Travel Times and Commuter Numbers (Long-Run Equilibrium)

	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
Dependent Variable:	$\log(\text{Commuter Numbers})$						
<i>Direct ICE</i>	6.236 (0.100)**	0.050 (0.006)**					
$\log(\text{Duration})$			-1.136 (0.008)**	-0.046 (0.007)**	-1.296 (0.056)**	-0.520 (0.069)**	-0.602 (0.121)**
$\log(\text{Pop Orig})$		0.195 (0.023)**	0.333 (0.003)**	0.169 (0.023)**	0.430 (0.006)**	0.257 (0.028)**	0.398 (0.121)**
$\log(\text{Pop Dest})$		0.876 (0.019)**	0.543 (0.003)**	0.908 (0.019)**	0.614 (0.006)**	0.941 (0.021)**	0.861 (0.097)**
$\log(\text{Distance})$			-0.466 (0.007)**		-0.321 (0.043)**		
<i>Year 2009</i>		0.033 (0.002)**	0.031 (0.005)**	0.034 (0.002)**	0.030 (0.006)**	0.022 (0.002)**	0.038 (0.009)**
R^2	0.03	0.13	0.74	0.13	0.58	0.05	0.20
N	124,284	124,250	123,680	123,680	123,680	89,440	2,068
Method	OLS	FE	OLS	FE	IV	FE, IV	PSM, IV, FE

* $p < 0.05$; ** $p < 0.01$

Table VII: Matching Regression

	Matching Regression, 1999	Comparison of Groups		
	ICE Introduction, 1999 - 2009	Treated	Untreated	t-statistic
$\log(\textit{Work Orig})$	0.151 (0.073)*	11.68	11.67	0.13
$\log(\textit{Work Dest})$	0.358 (0.076)**	11.98	11.99	-0.22
$\log(\textit{Pop Density Orig})$	0.235 (0.076)**	6.79	6.81	-0.31
$\log(\textit{Pop Density Dest})$	-0.611 (0.839)	7.00	7.04	-0.61
$\log(\textit{Area Orig})$	0.359 (0.071)**	5.83	5.80	0.43
$\log(\textit{Area Dest})$	0.344 (0.078)**	5.68	5.66	0.40
$\log(\textit{Distance})$	0.273 (0.051)**	5.28	5.29	-0.23
$\log(\textit{Duration})$	-0.769 (0.062)**	4.78	4.80	-0.26
$\log(\textit{Commuter Numbers})$	-0.067 (0.023)**	2.84	2.78	0.50
R^2	0.26			
N	34,505			

* $p < 0.05$; ** $p < 0.01$

Regression contains state-fixed effects and dummy variables for the region-type of each county within a pair.