# Motorways and firm performance: the case of Hungary

by

András Hann

Submitted to

Central European University

Department of Economics

In partial fulfillment of the requirements for the degree of

Master of Arts in Economics

Supervisor: Gábor Békés

Budapest, Hungary

2014

I, the undersigned [András Hann], candidate for the degree of Master of Arts in Economics at the Central European University Economics, declare herewith that the present thesis is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of work of others, and no part the thesis infringes on any person's or institution's copyright. I also declare that no part the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Budapest, 5 June 2014

Signature

# Abstract

I analyze the effect of motorways on firm performance in Hungary. I look at total factor productivity; the number of firms that open or close in a given municipality; and the firm's individual choice whether to cease operations by next year. I combine a database of 20,000 manufacturing firms in the 1992-2003 period with several sources of GIS data. Significant results are only obtained on pooled OLS data; fixed effects are only significant on a model restricted to municipalities with less than 10,000 inhabitants. When DiD models are estimated, I find no significant effects. I look at the number of new and closing firms using FE and FE Poisson methods – the latter of which fits count data better. These models do not yield significant results, either. Finally, I estimate a binary decision model of the firm whether to close down or not; results are insignificant as before.

# Acknowledgments

I would like to express my gratitude towards my supervisor, Gábor Békés for the valuable discussions and feedback he provided during the process of writing this thesis.

# Table of Contents

# List of Figures

# List of Tables

# Introduction

When Hungarian prime minister Gordon Bajnai spoke at the opening ceremony of Motorway M6/M60 in 2010, he said that "motorways are not just asphalt, guard rails, and a line of paint; they are the symbol of development and prosperity" (bama.hu, 2010).[1] Hungary has been steadily expanding its motorway network since 1989: it grew from 346 kilometers in total in 1989 to 1361 kilometers in 2013. Extensive investment in road infrastructure is not unique to Hungary; between 2000 and 2013 the European Union spent about €65 billion on cohesion funds targeted at road construction and renovation (European Court of Auditors, 2013). However, when these projects were audited, it was found that "the reporting on the achievements of the projects focused on the amount of funding used and physical output delivered, with almost no information about the effects of the projects on the local economy or actual gains for road users" (European Court of Auditors, 2013, p. 11).

Motorways may have a multitude of effects. First, they are proven to decrease total travel time and the number of accidents (European Court of Auditors, 2013). These are two clear benefits, but there are indirect effects as well: according to the literature on trade, transportation costs



[1]According to European Court of Auditors (2013, p. 5), a *motorway* is defined as "a road specially designed and built for motor traffic, which does not serve properties bordering it, and which: (i) is provided, except at special points or temporarily, with separate carriageways for the two directions of traffic, separated from each other by a dividing strip not intended for traffic or, exceptionally, by other means; (ii) does not cross at level with any road, railway or tramway track, or footpath; (iii) is specially signposted as a motorway." Whenever I use the term *motorway* in the paper, this is what I refer to.

determine the spatial structure of economic activities (Krugman, 1991); decreasing costs increase the volume of trade (Brakman, Garretsen and van Marrewijk, 2009); firm productivity is heavily influenced by market access (Fujita, Krugman and Venables, 1999). Looking at the micro-level, firms enjoy benefits of better market access and more potential customers – while they also have to face increasing competition, since others can now serve their customers. Through increasing returns to scale and the resulting agglomeration, firms can specialize and their productivity can increase.

In this thesis I look at the effect of new motorways on the local economy, more specifically the effect on the productivity, birth, and closure of manufacturing firms in Hungary. While effects of motorway constructions as early as the 1960's have been analyzed in the United States (Rephann and Isserman, 1994), early research yielded inconclusive results (Boarnet, 1995). However, more recent micro-level panel analyses have shown significant positive effects both on productivity (Holl, 2012, 2014; Lileeva and Trefler, 2010; Martín-Barroso, Núñez and Velázquez, 2013) and firm birth (Holl, 2004*a*,*b*). Hungarian literature on this topic is rare. While a few studies have been written recently (Németh, 2005; Ohnsorge-Szabó, 2006), these did not use the panel aspect of the data, nor did they account for possible endogeneity issues.

I compile a new dataset for Hungary from various sources which lets me estimate firm-level panel models of productivity and start and cease of operations – while keeping the aforementioned issues in mind. I also look at the firms' individual decision whether to shut down; I have found that this has not been analyzed before. My results (in contrast to previous Hungarian research) show only a very little effect if any.

In Chapter 1, I discuss previous literature and some methodological considerations for the analysis of the effect of highways. Chapter 2 introduces the various data sources that I use in the thesis. In Chapter 3 I analyze the effect of motorways on total factor productivity (TFP) using various econometric methods; Chapter 4 discusses the changes in the number of firms that open or close as well as the firm's decision whether to cease operations or not. Chapter 5 concludes.

# Chapter 1

# Literature review and methodological considerations

This chapter first presents previous results that analyzed the relationship between motorways and firms, then two methodological considerations that arise from the literature and are relevant for my analysis are discussed.

## 1.1 The effect of motorways

Motorways have a large effect on the accessibility of municipalities – the aforementioned trade literature (Krugman, 1991; Fujita, Krugman and Venables, 1999; Brakman, Garretsen and van Marrewijk, 2009) provides a theoretical background for such an analysis. They suggest that there are two main consequences of decreasing transportation costs (or, in other words, increasing market access): agglomeration of economic activity, and, in a somewhat related way, an increase in productivity. The relationship of these effects is widely analyzed in the literature (Ciccone and Hall, 1996; Davis and Weinstein, 2001; Brülhart, Crozet and Koenig, 2004; Brülhart and

Sbergami, 2009; Békés and Harasztosi, 2013) – I focus only on the latter of the two.

### 1.1.1 Productivity

Boarnet (1995) gives a review of early research on the effects of motorways and shows that up to that point, research was inconclusive, and if there was any effect, it is most likely just a shift of economic activity from other regions. Rephann and Isserman (1994) shows that there is only a small effect for rural counties, the real beneficiaries are more urbanized regions. This result is verified by Chandra and Thompson (2000) who use county-level US data and find differential effects across industries: some benefit more from decreasing transportation costs than others. Lileeva and Trefler (2010) analyze firms near the USA–Canada border; they find that better access to foreign markets raises productivity. Melo, Graham and Brage-Ardao (2013) conduct a meta-analysis of 33 studies. They find that when all relevant factors are properly considered, increasing public investment in roads by 10 percent increases output by only 0.5 percent.

Holl (2007) establishes a line of research with her exploratory analysis of the improvement of the Spanish motorways between 1980 and 2000, when the network was expanded from 1933 to more than 9000 kilometers in total length. She defines market access as

$$\mathrm{Acc}_i = \sum_j \frac{W_j}{c_{ij}^a} \tag{1.1}$$

The mass of location $j$ is denoted by $W_j$; this is simply weighted by dividing it with some cost measure of the distance between $i$ and $j$: $c^a$. The exponent $a$ is just an extra measure of friction as there may be difference between the costs depending on the goal of the trip.

This database is then used in Holl (2012) as well. Here the author first calculates total factor productivity (TFP) using a Cobb–Douglas production function (Holl, 2012, p. 5):

$$Y_{it} = A_{it} K_{it}^{\beta_1} K_{it}^{\beta_2} \tag{1.2}$$

where $A_{it}$, the TFP of the firm can be expanded into market accessibility and a firm level component (which is essentially a firm fixed effect and a random error, $V_{it} = \mathrm{e}^{\eta_i + u_{it}}$):

$$A_{it} = (\mathrm{MA}_{it})^{\delta} V_{it} \tag{1.3}$$

As it is commonly used in the literature, $Y$ is the value added, $K$ is capital use, $L$ is labor. $i$ refers to the firm, $t$ to the time period. Combining these and taking logarithms:

$$y_{it} = \beta_1 k_{it} + \beta_2 l_{it} + \delta \mathrm{ma}_{it} + \eta_i + u_{it} \tag{1.4}$$

To prevent errors resulting from the autocorrelation or endogeneity of TFP, the author uses methods described by Olley and Pakes (1996) to calculate it from the production function. Firm level productivity is then regressed on market potential using several instrumental variables (historical road network, terrain ruggedness) to control for the possible endogeneity in the location of motorways (they could be built in regions that are already growing more or expected to grow). The estimations show a significant positive effect.

Further analysis of Spanish data is presented in Holl (2014): the effects are estimated comparing urban and rural areas – again, on the firm-level database and instrumental variables using the same idea (historical roads, geographical features) as before. The results confirm that of Rephann and Isserman (1994): rural areas benefit less than already urbanized ones.

Martín-Barroso, Núñez and Velázquez (2013) follow a very similar pattern of estimating TFP from the production function then using this as the dependent variable. The focus of their analysis is the different aspects of accessibility: it is more important for firms to have good access to commodities than to workers.

Ghani, Goswami and Kerr (2012) take a different approach. The subject of their analysis is the Golden Quadrilateral project, which vastly improved road connections between the four largest cities in India. Instead of using TFP, simpler measures are explained using the models:

the number of new plants, employment, and output. The authors use a difference-in-difference method, which compares the average values of treated and non-treated subjects before and after some intervention (building a highway).[1] The results are not in line with Holl (2014): the authors show that there is a significant positive effect in rural areas as well. The difference-in-difference method is also used by Datta (2012), who looks at another aspect of firm behavior after the Golden Quadrilateral was built: he finds that firms decrease their inventory because of the more reliable supply through better roads.

### 1.1.2 Firm birth

Holl (2004*a*) analyzes how motorways (and accessibility in a broader sense) influences the number of new firms in Spanish municipalities at a sectoral level. The number of new firms, according to the author's model, is affect by a variety of factors – she calculates measures of inter- and intra-regional demand accessibility and supplier accessibility; population, wages, and labor force qualification are controlled for as well. The distance to motorways is included directly. This means that there is a *ceteris pariubs* effect of motorways: when the accessible market is of the same size, better motorway access still attracts more firms – the effect is significant in all sectors but chemical products. Holl (2004*b*) conducts a very similar analysis for Portugal and finds that there is a strong and significant preference of firms for locations not farther than 10 kilometers from the motorway.

### 1.1.3 Hungarian results

The effects of motorways on economic activity has not been widely analyzed in Hungary. One example, a broad, descriptive overview is the paper by Németh (2005). The author analyzes the relationship between proximity of municipalities to motorways and unemployment or income

[1] I use this method in Section 3.3, where I describe it in detail.

per person in the said municipality. While he controls for the education level of the population as well as the distance to the western border, a cross-sectional OLS analysis cannot be interpreted as more than mere correlations. As a "proto-difference-in-differences" method, the author visually compares trends of variables in question near to and farther away from motorways. Ohnsorge-Szabó (2006) takes a similar approach of estimating the effect (also on municipality-level employment and income) using cross-sectional OLS – which suffers from the same problems. It should be noted, though, that he includes railway accessibility as another measure, which is an idea worth investigating. On the other hand, railway investments (and thus improvements in railway accessibility) are close to nonexistent in Hungary, so such an effect is not easy to measure properly.

As far as I know, firm-level analysis of the effect of motorways in Hungary does not exist. One paper uses micro-level data, but analyzes a different question: Márk (2013) estimates the effect of a new motorway on property prices. The author uses a difference-in-difference estimation to assess the effect of the M6/M60 motorway that was opened in 2010 – since there is a clear one-time intervention and intertemporal changes are used for identification, this is likely the most econometrically sound paper on Hungarian motorways.

## 1.2 Methodological considerations

In this section I discuss methodological considerations that are relevant throughout the paper; estimation-specific concepts are discussed when they arise.

### 1.2.1 Endogeneity of highway construction

A key issue in my analysis is whether motorways exogenously affect TFP or there is an inverse causality as well – or maybe there is a third variable that simultaneously affects both. Melo, Graham and Brage-Ardao (2013) note that this is a frequent problem in analyses of the effect of motorways. It is reasonable to believe that motorways are built in locations that are already developing. Chandra and Thompson (2000) tested this hypothesis for the United States and rejected inverse causality. Bronzini and Piselli (2009) show that in Italy there is Granger causality between infrastructure investments and productivity growth (i.e. the first happens before the other). Pradhan and Bagchi (2013), however, find signs of inverse causality in India. Holl (2004*a*) also tests whether growth affects motorway construction – she finds no significant effect. She explains that fixed effect estimates are consistent in her case because Spanish motorways are built very close to existing main roads, so municipalities are already either well-connected or not. In this case, this characteristic is removed when the fixed effects methods are used, so only the differential effect of new motorways is included. It is thus possible that Western results found no inverse causality because of the already well-developed infrastructure (where motorways improve existing connections but do not create new ones) – while in developing India, motorway construction plans are still influenced by current economic events and future expectations.

However, in some papers, there are instrumental techniques used, and in several cases, the instrument is the historical main road network (Holl, 2012, 2014). In the case of Hungary, Österreichisches Staatsarchiv (2014) has published the Second Military Survey of the Habsburg Empire online. Instead of digitizing the whole database, I make two observations: 1) motorways run along primary roads (see Figure 2.3); and 2) by looking at the Second Military Survey map, we can confirm that the location of the primary roads in the past is in many cases the exact same today as in 1869 – and when this is not the case, the difference is only a few kilometers. Hence, the old road network can be very well proxied by the new road network, the distance from which, in turn, could be used to instrument the distance to motorways.

There is still an unresolved issue, though: according to Holl (2014), in order to instrument panel models, one needs an instrument that varies over time. Unfortunately, my candidate for an instrument is not like this. A cross-sectional analysis, even if the distance to the motorways is instrumented, would omit the firm-level fixed effects. This would in the end still lead to biased estimates. Thus, I do not use an IV method but assume that motorway construction is exogenous.[2] A small trick that I implement, though, is restricting the sample to firms in municipalities with fewer than $10,000$ inhabitants: even if transport infrastructure is placed nonrandomly, it is usually designed to connect larger cities (Hornung, 2012); the treatment of small settlements can thus be considered random.

### 1.2.2 The functional form of the effect of motorways

There is no solid reason to believe that the relationship between the distance to the motorway and TFP of a firm or the number of closing or opening firms is linear (even if the log-log specification fits better than the log-linear or linear-linear). The effect could steeply decrease then level off or it could be not significant very close to the motorway because of the possible disadvantages (noise, higher level of traffic) but significant somewhat farther. With cross-sectional data, especially in the case of hedonic property price analyses, semi-parametric models are commonly used to estimate such relationships (see Koster and Rouwendal (2012) or Bontemps, Simioni and Surry (2008)). In these cases, the model is not imposed to be linear but estimated as a partially linear and nonparametric model:

$$y = \beta' x + m(z) + \eta \tag{1.5}$$

where $m(.)$ is estimated as a locally weighted polynomial regression (Bontemps, Simioni and Surry, 2008). However, Lokshin (2006) notes that these models do not work directly with panel data, so I resort to using log-log models and models with distance-band dummies, which indicate whether the firm is up to 20 km, 20–40 km, or more away from the highway – this gives some flexibility to the estimation.

[2]Hungary is already a developed country, and just as in Spain, motorways run rather close to already existing main roads (see Figure 2.3). I do not have observations from enough time period to be able to test this hypothesis directly.

# Chapter 2

# Data and descriptives

## 2.1 Firms

The firm-level database that I used is a stripped-down version of the IEHAS-CEFiG database, which I received from the authors of Békés and Harasztosi (2013). It includes data on all Hungarian manufacturing firms from NACE European Classification of Economic Activities categories 17–37 between 1992 and 2003 that employ at least five people. The dataset is described in detail by Békés, Muraközy and Harasztosi (2011). This is an unbalanced panel dataset. Wooldridge (2012) warns that estimations using unbalanced panel data may be biased do to unknown mechanisms of dropping out from the sample. However, since this is not really a *sample* but the full population, we know that if a firm leaves the database it is because it dropped below 5 employees (and usually went out of business), so this is not a problem. The dataset contains observations for $20,867$ firms in $1,425$ settlements in Hungary. Note that this means that there are many settlements with no manufacturing firms, since in total there are about $3,150$ settlements. For each observation, there is a firm ID, year, settlement, microregion ID, county ID, sector ID, number of workers, labor force in the area as a measure of agglomeration, and a dummy whether the firm trades. Békés and Harasztosi (2013) also used this dataset to calculate TFP for each firm in each year.[1] These results were included in the data I received from the authors. For confidentiality reasons, however, I was not given access to firms' income, factor use, or other sensitive data.

Table 2.1 summarizes the number of firms in each year by sector. We can see that while the database does not cover the entire production sector, the sectors that are covered are actually rather diverse. The differences in the typical proportion of each sector in a given year are great but not surprising: only a few firms deal with petroleum products or nuclear fuels, while metal product and machinery firms are numerous. Also, a general increasing trend in the numbers is apparent.

Figure 2.1 shows the spatial distribution of firms in the sample in 2002. It is immediately visible that Budapest is overrepresented; more than one fourth of all firms are in the capital. Otherwise, there is no apparent spatial pattern: while the southeastern part of the country may seem to have more firms, the larger area of these settlements can be deceiving.

## 2.2 Road network

The source of road network data is the OpenStreetMap database (osm.org, 2014).[2]. Using the OverPass Turbo interface[3] of the OverPass API[4] I downloaded the network of motorways, trunk roads, primary, secondary, and tertiary roads as they were between 1992 and 2012 – this way I could follow the development of the motorway network. Then, using the GraphHopper route planner engine and its API[5] I calculated the distance and duration of the fastest trip from each

[1]The authors use a method similar to that of Holl (2012), discussed in more detail in Section 1.1. Thus, the TFP is correctly estimated without endogeneity/omitted variable bias problems.

[2]Regarding data reliability, it may be a concern that the map can be edited by anyone and the changes are immediately implemented in the database. However, random checks of consistency with commercially produced Google Maps has shown that some areas may be less detailed (e.g. fewer small roads), but inaccuracies are extremely rare, and the main road network of Hungary fully exists in the database

[3]`http://overpass-turbo.eu/`

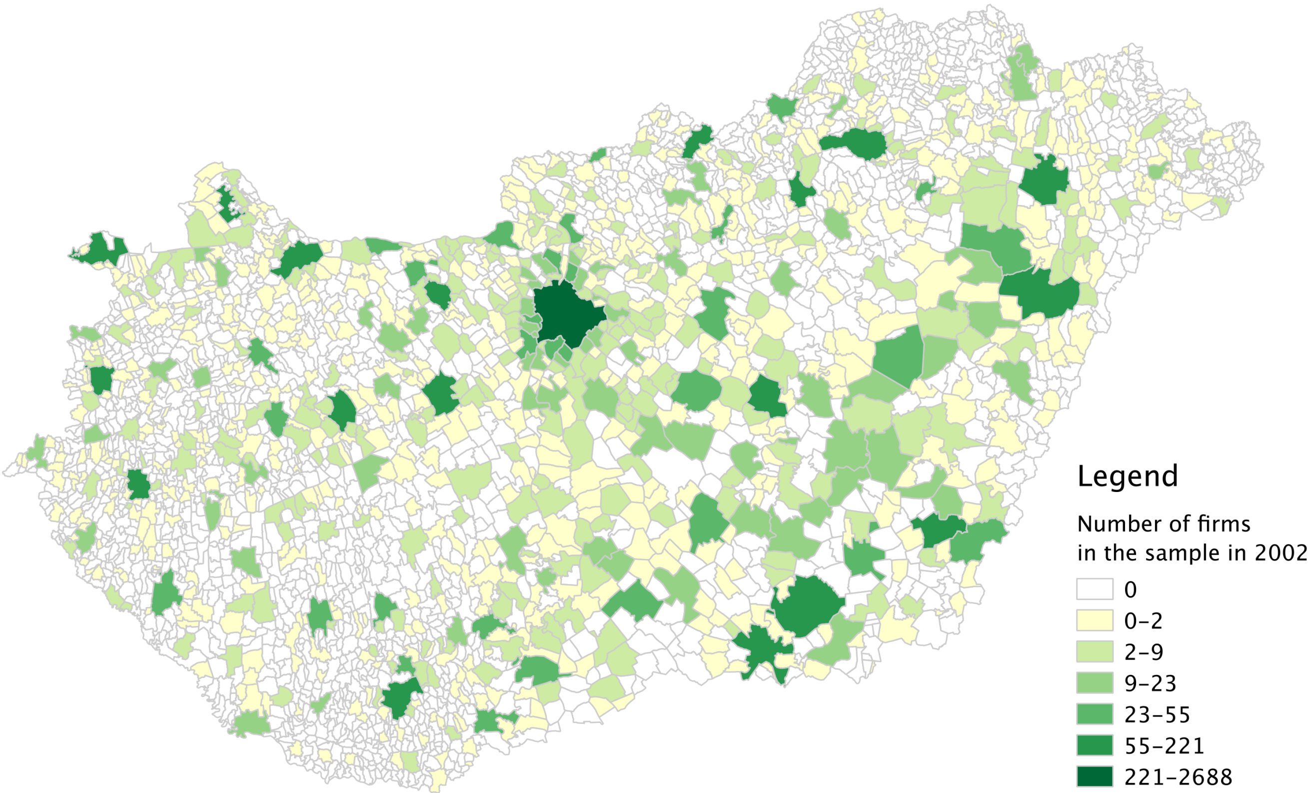
[4]`http://overpass-api.de`

[5]`http://graphhopper.com/`

**Table 2.1: Yearly number of firms by sector.**

| | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Textiles | 265 | 295 | 317 | 335 | 382 | 454 | 488 | 487 | 469 | 460 | 441 | 428 |
| Wearing apparel; dressing and dyeing of fur | 505 | 567 | 568 | 580 | 614 | 684 | 748 | 778 | 742 | 728 | 685 | 635 |
| Tanning and dressing of leather; luggage, handbags… | 187 | 214 | 208 | 226 | 238 | 267 | 275 | 274 | 259 | 257 | 245 | 210 |
| Wood and of products of wood and cork, except furniture… | 342 | 415 | 456 | 472 | 530 | 609 | 668 | 726 | 687 | 668 | 684 | 658 |
| Pulp, paper and paper products | 73 | 89 | 91 | 104 | 113 | 127 | 142 | 148 | 157 | 162 | 168 | 169 |
| Publishing, printing and reproduction of recorded media | 614 | 687 | 734 | 741 | 781 | 865 | 926 | 960 | 779 | 774 | 781 | 769 |
| Coke, refined petroleum products and nuclear fuel | 6 | 7 | 7 | 6 | 5 | 5 | 6 | 5 | 4 | 4 | 5 | 6 |
| Rubber and plastic products | 351 | 382 | 422 | 452 | 480 | 557 | 630 | 670 | 701 | 696 | 729 | 718 |
| Other non-metallic mineral products | 283 | 301 | 331 | 357 | 392 | 433 | 447 | 473 | 450 | 452 | 455 | 442 |
| Basic metals | 89 | 101 | 110 | 127 | 126 | 127 | 141 | 141 | 139 | 138 | 141 | 132 |
| Fabricated metal products, except machinery and equipment | 844 | 959 | 1040 | 1144 | 1246 | 1481 | 1598 | 1711 | 1618 | 1612 | 1597 | 1624 |
| Machinery and equipment n.e.c. | 873 | 952 | 1014 | 1074 | 1153 | 1253 | 1292 | 1313 | 1190 | 1172 | 1169 | 1135 |
| Office machinery and computers | 54 | 60 | 62 | 69 | 66 | 80 | 82 | 82 | 78 | 78 | 78 | 72 |
| Electrical machinery and apparatus n.e.c. | 251 | 286 | 309 | 330 | 370 | 413 | 431 | 455 | 423 | 419 | 418 | 422 |
| Radio, television and communication equipment and apparatus | 155 | 192 | 211 | 226 | 243 | 265 | 282 | 295 | 294 | 286 | 284 | 283 |
| Medical, precision and optical instruments, watches and clocks | 227 | 269 | 294 | 302 | 315 | 369 | 411 | 426 | 378 | 375 | 387 | 381 |
| Motor vehicles, trailers and semi-trailers | 81 | 102 | 123 | 124 | 130 | 157 | 170 | 172 | 180 | 183 | 186 | 172 |
| Other transport equipment | 35 | 41 | 41 | 44 | 56 | 69 | 77 | 82 | 79 | 82 | 82 | 86 |
| Furniture; manufacturing n.e.c. | 320 | 369 | 382 | 406 | 470 | 551 | 614 | 618 | 610 | 605 | 601 | 595 |
| Recycling | 14 | 20 | 26 | 30 | 35 | 41 | 47 | 55 | 53 | 51 | 57 | 57 |
| *Total* | 5569 | 6308 | 6746 | 7149 | 7745 | 8807 | 9475 | 9871 | 9290 | 9202 | 9193 | 8994 |

**Figure 2.1: Spatial distribution of firms in the sample in 2002. Source: own work based on CEFiG and OpenStreetMap data.**




settlement to Budapest as well as the closest important border crossing.[6] I decided to analyze the fastest trip because in many cases, the motorway is somewhat longer in distance but shorter in time – thus, if I looked at distance, I may have ignored motorways completely.



[6]I classified border crossings "important" if they lay on roads that are members of the International E-road Network. To be precise, I calculated the trip duration and distance to the nearest settlement to the border on the E-roads, which are Ártánd, Hegyeshalom, Hercegszántó, Letenye, Nagylak, Parassapuszta, Rajka, Rédics, Röszke, Szentgotthárd, Tiszabecs, Tornyiszentmiklós, Tornyosnémeti, Udvar, Vámosszabadi, and Záhony.

## 2.3 Motorway opening

I use the dataset compiled by Madura (2013), who told me in an email that he compiled this database from news articles and other reputable resources he found, predominantly the website of the National Infrastructure Developing Private Company Limited (NIF Zrt.).[7] However, a detailed reference list was not available.[8] I then used this data to amend the existing OpenStreetMap data then calculate routes using this updated database.

Travel time data, which I derived from these sources, confirms that the construction of motorways was not only significant in the total length of the network, but also heavily affected the duration of a trip to Budapest from other municipalities. Figure 2.2 shows the change in the kernel density function (a nonparametric probability density function) of travel time to Budapest: shrinking from 144 to 129 minutes, the average trip is 10.4 percent shorter; however, this is not weighted by the number of actual trips, which would probably show an even more significant reduction in travel time.

The market access measure calculated by Holl (2007) is adequate in Spain, since it is a relatively large but compact geographical region bordered by mountains and seas. However, in the case of Hungary I do not use this measure: Hungary is a small open economy and because of this it may not be sensible to cut off relevant territories at the border. While my analysis ends before Hungary's accession to the EU, this problem became even more complicated since then: there are borders with no control at all (Schengen countries), EU borders with border control, and external borders. Trade spills over open borders more than closed one (Lileeva and Trefler, 2010); these foreign municipalities would also have to be included in such a market access measure. Instead of trying to set some weights and obtaining the required detailed foreign data, it is more feasible to just calculate the distance to the motorway and use it as a market



[7] http://nif.hu

[8] He has also uploaded his work to Wikipedia, where it is accessible at https://hu.wikipedia.org/wiki/Aut%C3%B3p%C3%A1ly%C3%A1k_Magyarorsz%C3%A1gon

access measure.

**Figure 2.2: Kernel density function of travel time to Budapest in 1992 and 2012. Source: own work based on OpenStreetMap data.**

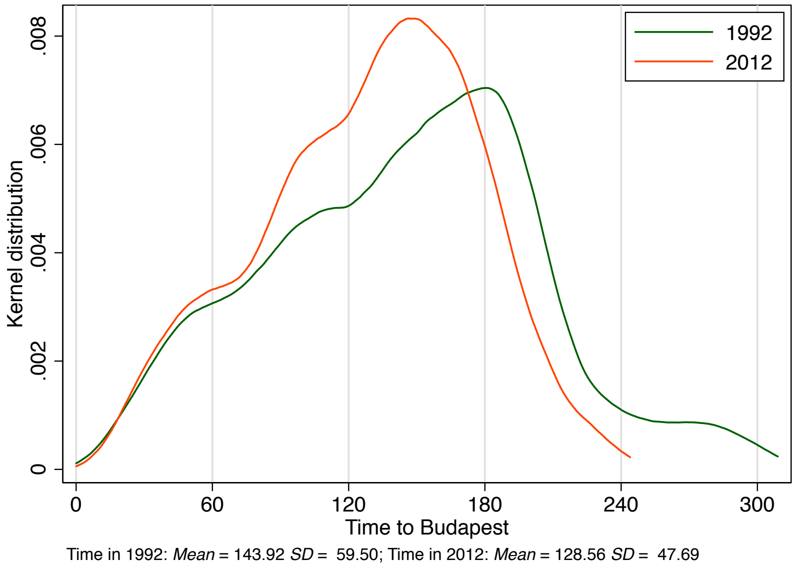
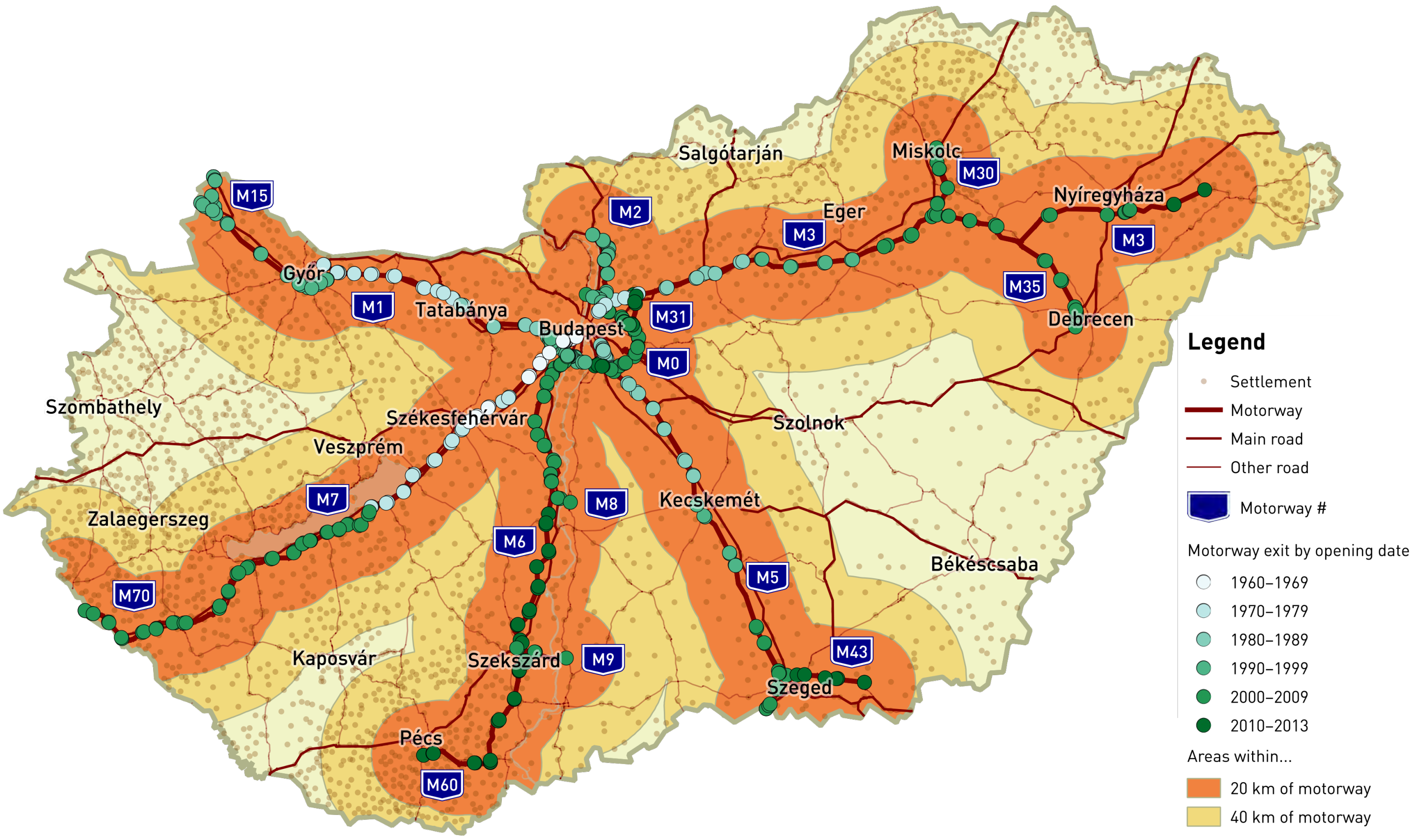



Figure 2.3 shows a map of main roads in Hungary and motorway exits and their opening dates; I made this map using the previously mentioned data sources. Figure B.1 in the Appendix shows the changes that happened between 1992 and 2012 in the required time to reach the municipalities of Hungary from Budapest by car.

In the period between 1992–2003, for which there is data on firms, motorway construction was not as fast as after that; about 300 kilometers of motorways were built in total.[9] In 1994 and 1996, motorway M1 was extended from Győr to the Austrian border and a small portion of M0, the ring road around Budapest was built. In 1997, M5 was built between the north of Kecskemét to the south of the city. 1998 brought the extension of M3 between Gyöngyös and Füzesabony, as well as the continuation of M5 to Kiskunfélegyháza. In 1999 M2, a road that does not really fit the definition of a motorway, but is often considered as such, was opened between Budapest

[9] This paragraph is built on the data from Madura (2013).

**Figure 2.3: Map of important roads in Hungary including motorways and motorway exits. Source: own work based on OpenStreetMap data.**




and Vác. In 2002, M3 between Füzesapony and Polgár, M30 between Igrici and Emőd, M7 between Balatonaliga and Zamárdi was finished. Finally, in 2003 a new bridge on the Danube near Szekszárd was opened as the only segment of the planned M9 motorway; M30 was extended from Emőd to Nyékládháza. These developments decreased the travel time to Budapest by more than 30 minutes for 272 settlements, which is about 8 percent of all municipalities in the country – this is clearly a remarkable improvement. The length of new motorways in each year as well as the total length of the Hungarian motorway network is displayed in Table 2.2.

**Table 2.2: History of Hungarian motorways in the analyzed period. Source:Madura (2013).**

| Year | New motorways (km) | Total motorways (km) |
|---|---|---|
| 1994 | 37 | 398 |
| 1996 | 42 | 440 |
| 1997 | 16 | 456 |
| 1998 | 77 | 533 |
| 1999 | 38 | 571 |
| 2002 | 67 | 638 |
| 2003 | 29 | 667 |

## 2.4 Municipalities

The source of municipality-level data is the Hungarian Central Statistical Office's Statistical Database of Municipalities (KSH T-STAR). This database contains several hundred observed characteristics of 3166 settlements since 1990[10]. I use one data series throughout my thesis: the population of the municipality in the given year.



[10] 3166 is not the actual number of settlements as there are some cases where settlements split up (generally small neighborhoods becoming independent); in this case starting the next year, the new settlements are in the database, while older data cannot be broken up and thus these municipalities will still be included

# Chapter 3

# Productivity

In this chapter I analyze firm productivity changes in response to better access to the motorway network. First, I show simple OLS results, which I expand into fixed effects estimations, which control for time-invariant characteristics of firms as well as general trends over time. Finally, I look at a difference-in-differences approach, which compares the changes over time between groups that received a treatment and groups that did not.

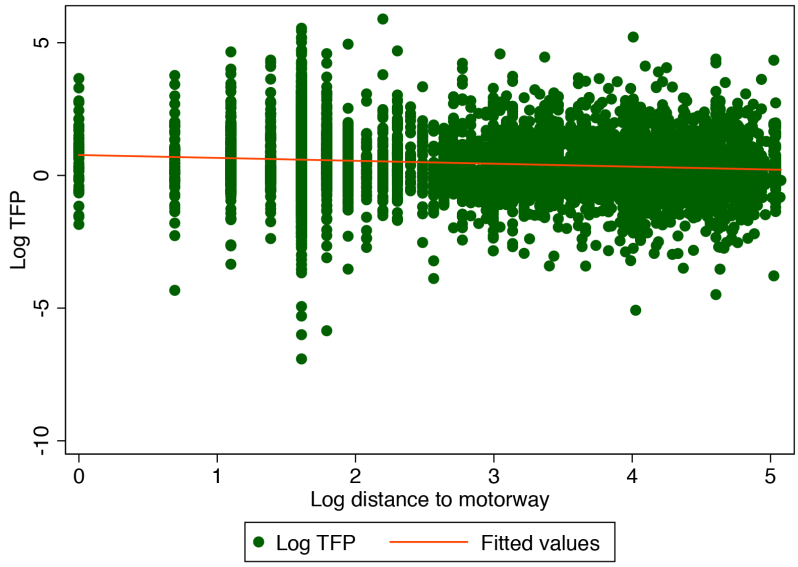
## 3.1 First look: OLS estimations

The first and simplest way to look at the connection between the distance to the motorway and TFP is a cross-sectional regression run on a sample restricted to one year. This simple model can be estimated as

$$\log(TFP)_{i,2003} = \beta_0 + \beta_1 \log(dst)_{i,2003} + u_{i,2003} \tag{3.1}$$

where $\log[(TFP)_{i,2003}$ is the log TFP of firm $i$ in 2003; $\log(dst)_{i,2003}$ is the log distance of firm $i$ from the motorway in 2003. There are several reasons to choose a log-log functional form:

**Figure 3.1: Correspondence between log distance to motorway and log TFP in 2003. Source: own work based on CEFiG data.**




first, a straight line fits such a relationship better.[1] Also, interpretation is easier as in this case the coefficient is a simple elasticity: a one percent increase in $dst_{i,2003}$ leads to a $\beta_1$ percent increase in $TFP_{i,2003}$.

In this model, my null hypothesis is that the relationship is negative: TFP of firms that are farther away from the motorway is likely lower than those that are close-by. The results of estimating (3.1) are shown in (1) of Table 3.1 and also graphically in Figure 3.1; these confirm the hypothesis: there is a significant negative relationship between the variables. However, this model clearly suffers from an omitted variable bias as we did not control for several possible factors that may affect TFP. For example, less developed parts of the country may be less productive – in the case of Hungary, this is the Eastern part of the country. This part also has worse access to Western Europe: even if two firms are the same distance from the motorway, cargo has

[1]A one-variable regression model yields $\overline{R}^2 = 0.017$ in the case of the log-log model while only $\overline{R}^2 = 0.005$ for the linear-linear model.

**Table 3.1: Cross-sectional models of relationship between distance to motorways and productivity.**

| | Log TFP | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) OLS | (2) OLS | (3) OLS | (4) OLS | (5) OLS | (6) OLS | (7) OLS |
| Log distance to motorway | -0.0986*** | -0.0812*** | -0.102*** | -0.101*** | -0.0508*** | -0.0579*** | -0.0293** |
| | (0.0116) | (0.0121) | (0.0105) | (0.0104) | (0.0142) | (0.0145) | (0.0130) |
| Eastern Hungary | | -0.130*** | -0.0289 | -0.0109 | -0.0162 | -0.0118 | -0.0309 |
| | | (0.0265) | (0.0244) | (0.0242) | (0.0242) | (0.0243) | (0.0218) |
| Log number of employees | | | 0.260*** | 0.206*** | 0.206*** | 0.207*** | 0.238*** |
| | | | (0.00986) | (0.0111) | (0.0111) | (0.0111) | (0.0108) |
| Trader | | | | 0.321*** | 0.629*** | 0.390** | 0.380** |
| | | | | (0.0258) | (0.0753) | (0.171) | (0.154) |
| Log distance to motorway × trader | | | | | -0.0842*** | -0.0798*** | -0.0547*** |
| | | | | | (0.0192) | (0.0194) | (0.0178) |
| Log time to nearest border crossing | | | | | | -0.0890*** | -0.0747*** |
| | | | | | | (0.0268) | (0.0239) |
| Log time to nearest border crossing × trader | | | | | | 0.0547 | 0.0277 |
| | | | | | | (0.0356) | (0.0319) |
| Sectoral dummies | N | N | N | N | N | N | Y |
| Constant | Y | Y | Y | Y | Y | Y | Y |
| Observations | 6,148 | 6,148 | 6,167 | 6,167 | 6,167 | 6,157 | 6,157 |
| R-squared | 0.013 | 0.017 | 0.123 | 0.145 | 0.148 | 0.150 | 0.304 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Note*: Sample restricted to 2003. Distance to motorway is calculated on the municipality level. "Eastern Hungary" indicates that the firm is in one of the following regions: Northern Hungary, Northern Great Plain, Southern Great Plain.

to travel farther from here. By adding this variable in (2) of Table 3.1 we see that there is indeed a significant effect and the effect of distance decreases.

Still many possibly relevant factors were not included, which are added one by one in (3) to (6) of Table 3.1: the log number of employees, whether a firm trades or not, and the logarithm of the time it takes to reach the nearest border crossing. Trading is interacted with the latter one as well as the distance to the motorway: it is possible that accessibility affects trading firms differently than non-traders. Finally, in (7), sectoral dummies are included as well, arriving at the following model:

$$\begin{aligned}\log(TFP)_{i,2003} = \beta_0 + \beta_1 f(dst_{i,2003}) + \delta_1 easthun_i + \beta_2 logempl_{i,2003} + \delta_3 trader \\ + \delta_4 trader \times \log(dst)_{i,2003} + \beta_3 \log(bordertime)_{i,2003} \\ + \delta_5 trader \times \log(bordertime)_{i,203} + \delta_i sector + u_{i,2003}\end{aligned} \tag{3.2}$$

Function $f(\cdot)$ accounts for the flexible functional form mentioned in Section 1.2.2, it can be either simply $\log(x)$ or two dummies for the first 20-kilometer bands (where a third dummy, being farther than 40 kms is the baseline group.)

Adding the number of employees turns the Eastern Hungary indicator insignificant. This means that it is not whether the firm is in Eastern Hungary or not that has an effect, it is rather the size of the firm; companies of different size are not evenly distributed across the country. Further variables do not change the significance of others, the newly added ones seem to subtract from the effect of distance, which means that firms in different distance differ in other aspects as well.

There are still other factors that are not included, many unobserved – and unobservable – aspects of the firms and location and time fixed effects are not included. Thus, it has to be stressed that these results do not, in fact, show any causality whatsoever; we have just observed some *correlation* between the variables. The next sections aim to fix these issues.

## 3.2 Fixed effects

Simply adding year dummies to the previous model and pooling the observations across time periods would clearly improve the previous estimations. However, it would still not solve the problem of time-invariant characteristics of firm, i.e., firm *fixed effects*. In this section I discuss the various methods to account for these characteristics then estimate the effect of motorways on productivity using said methods.

### 3.2.1 Modeling fixed effects

A general model with effects can be written as:

$$y_{it} = \beta_0 + \beta' x_{it} + a_i + u_{it} \tag{3.3}$$

where $a_i$ stands for the individual-specific fixed effect. Wooldridge (2012, p. 461, p. 484) shows two types of methods of incorporating this effect into the model. The first one is taking differences across time periods (with period dummies), in which case

$$\begin{aligned} y_{it} - y_{it-1} &= \delta' year + \beta' x_{it} - \beta' x_{it-1} + a_i - a_i + u_{it} - u_{it-1} \\ \Delta y_{it} &= \delta' year + \beta' \Delta x_{it} + \Delta u_{it} \end{aligned} \tag{3.4}$$

which eliminates the individual fixed effect and thus the bias resulting from any such variables that would have been omitted. Another approach – "proper" fixed effects – is to time-demean the data, i.e. subtract the average value of the variable over time for the given individual:

$$\begin{aligned} y_{it} - \bar{y}_i &= \beta_0 - \beta_0 + \beta'(x_{it} - \bar{x}_i) + a_i - a_i + u_{it} - u_i \\ &= \beta'(x_{it} - \bar{x}_i) + u_{it} - u_i \end{aligned} \tag{3.5}$$

which, again, removes the individual fixed effects. Wooldridge (2012) notes that we can simply estimate this model using dummy variables for each individual, which is simpler to do and yields the same results:

$$y_i t = \beta_0 + \beta' x_{it} + \delta_i + u_{it} \tag{3.6}$$

Note that in this case a constant term is also included. The author notes, however, that it is the best to use the built-in commands of statistical software, which take care of these issues.

However, removing the fixed effect presents further issues: we cannot estimate the effect of variables that are constant over time for a given individual. When there are small year-to-year changes but possibly large changes over the observed period, a time-demeaned or dummy variable regression is likely to perform better as there is more variance in the explanatory variable. In my sample there are $77,956$ observations for which a change in distance to the motorway can be calculated, however, only $5,043$ of these are actual nonzero changes. On the other hand, in Figure 2.2 and the accompanying text I have shown that there has been a large change in distances to the motorway between 1992 and 2012. Firm data is only available until 2003, in this period the average distance of firms to the motorway shrunk from 50.55 km (sd 57.42) to 34.41 km (sd 36.65). This suggests that a fixed effect, not a first difference estimation is optimal in this case. Also, if the effect did not unfold in one year, FD estimations may be incorrect. Hence, I present the results of fixed effect estimations here; differenced model results are shown in the appendix, Table A.1 and A.2.

### 3.2.2 Estimations with all sectors in the sample

I estimate variations of the following model:

$$\log(TFP)_{it} = \beta_0 + \beta_1 f(dst_{it}) + \beta' x_{it} + \delta_i + \delta_t + u_{it} \tag{3.7}$$

where $\log(TFP)_{it}$ is the log TFP of firm $i$ in year $t$, $f(\cdot)$ is a function of the distance of firm $i$ to the motorway in time period $t$,[2] $x_{it}$ are other control variables (log number of employees, trader dummy, log distance to motorway $\times$ trader dummy, trader $\times$ log time to nearest border crossing), $\delta_i$ is a firm fixed effect, and $\delta_t$ is a year fixed effect. $f(\cdot)$ is the same as before: either the logarithm of the variable or a dummy of 0–20 or 20–40 kilometers, where 40+ is the baseline.

Fixed effect estimation results are shown in Table 3.2. Column (1), (2), and (3) are simple

[2] I could also use the trip duration to Budapest. However, the correlation between the two is very high, about 80 percent, and the intervention by the goverment is motorway construction – even if it then decreases the time required to reach Budapest. Hence I stick to this measure of accessibility throughout the thesis.

**Table 3.2: Fixed effects models of relationship between distance to motorways and productivity.**

| | Log TFP | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) OLS | (2) OLS | (3) OLS | (4) FE | (5) FE | (6) FE | (7) FE | (8) FE |
| Log distance to motorway | -0.0625*** | | | 0.00835 | -0.00311 | -0.00158 | | |
| | (0.00447) | | | (0.0116) | (0.0106) | (0.0113) | | |
| *Distance to motorway...* | | | | | | | | |
| 0–19 km | | 0.1543*** | 0.12*** | | | | 0.0479 | 0.108** |
| | | (0.0122) | (0.0209) | | | | (0.0297) | (0.0515) |
| 20–39 km | | 0.138*** | 0.1894*** | | | | -0.000694 | 0.00207 |
| | | (0.011) | (0.018) | | | | (0.0223) | (0.0419) |
| Year fixed effect | Y | Y | Y | Y | N | N | N | N |
| Year × sector fixed effect | N | N | N | N | Y | Y | Y | Y |
| Other controls | Y | Y | Y | N | N | Y | Y | Y |
| Constant | Y | Y | Y | Y | Y | Y | Y | Y |
| $F$-test of distance dummies | | 113.69*** | 71.24*** | | | | 1.58 | 2.57* |
| p-value | | 0 | 0 | | | | 0.2065 | 0.0764 |
| Observations | 63,854 | 63,854 | 63,854 | 63,961 | 63,961 | 63,854 | 63,854 | 19,788 |
| $R^2$ (within) | 0.190 | 0.191 | 0.227 | 0.046 | 0.089 | 0.107 | 0.107 | 0.131 |
| Sample | Full | Full | Pop < 10000 | Full | Full | Full | Full | Pop < 10000 |
| Firms in sample | | | | 12,020 | 12,020 | 11,999 | 11,999 | 3,802 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Note*: “Other controls” are: log number of employees, trader dummy, log distance to motorway × trader dummy, trader × log time to nearest border crossing. Distance to motorway is calculated on the municipality level. In (2), (3), (7), and (8) the baseline distance group is 40+ km. OLS results are different from those in Table 3.1 because of the different explanatory variables. Budapest is omitted from the sample.

pooled OLS models. In such a model, all observations are put together, including only time fixed effects (i.e. dummies for each year) but not individual fixed effects:

$$\log(TFP)_{it} = \beta_0 + \beta_1 f(dst_{it}) + \beta' x_{it} + \delta_t + u_{it} \tag{3.8}$$

(1) estimates the effects in a log-log specification, (2) uses distance dummies, and to remedy the selection bias (townships non-randomly receiving the treatment of motorways), (3) is run on a sample restricted to municipalities of fewer than $10,000$ inhabitants. All three models

yield a significant and negative effect of distance: being closer to the motorway increases TFP. According to (1), for example, halving the distance to the motorway (i.e. decreasing it by 50%) leads to a $50 \times 0.0624084 = 3.12042\%$ increase in TFP. While this is statistically significant, it is rather small in the economic sense. Also, these models are based on an extension of the model in Equation (3.2), but they still do not control for firm fixed effects, which is shown in Columns (4)–(8).

Columns (4)–(8) show various versions of the estimation of Equation (3.7), which all include firm fixed effects. (4) is a model of logarithmic distance; it includes year fixed effects and no further controls. The coefficient of distance becomes insignificant at any usual significance level. This shows that there was indeed an omitted variable bias: time-invariant firm characteristics, such as the quality of management, not the distance to the motorway is what affects TFP, which means that the relationship that we see in one-period or pooled OLS models was just a correlation and firms with higher TFP received better access to motorways.

In (5), there are still no other control variables, however, instead of pure year effects, there are interactions of year and sectoral dummies: it could be the case that there are some sectors in the sample that generally prosper while others were in a recession, and not taking this into account may lead to many different types of errors based on the structure of the omitted trend. The effect is once again insignificant. When other control variables (log number of employees, trader dummy, log distance to motorway $\times$ trader dummy, trader $\times$ log time to nearest border crossing) are added in (6), the point estimate is even lower in absolute value and is insignificant. Even if we assume that the estimate is unbiased and thus with more observations we would achieve smaller standard errors, the effect is absolutely negligible in the economic sense.

(7) uses dummies for categories of distance to the motorway along with time and firm fixed effects as well as the aforementioned further control variables. Once again, we do not see a significant effect. In (8) the same specification is run but this time on a sample restricted to settlements under $10,000$ inhabitants as in (3). This time, the effect is significant (both statistically

and economically): those firms that are in small towns not farther than 20 kilometers from the motorway are 10.8% more productive than those that are farther than 40 kilometers. This result can stem from two factors: a) I have successfully remedied the problem of nonrandom selection, or b) better motorway access has much higher benefits for small firms in small municipalities, as it is a relatively larger increase in market access.

### 3.2.3 Sector-by-sector estimations

It can easily be the case that some sectors differ from others in the effect of being close to motorways: more transportation-intensive sectors can benefit more than others. For example, if the produce is bulky and it is typically transported over a long distance (for example heavy machinery, which is exported to foreign countries), the availability of easy transportation is more important than for, for example, weaved baskets that are light and sold locally. I estimate Equation (3.7) separately for each of the sectors:

$$\log(TFP)_{it}|sector = \beta_0 + \beta_1 f(dst_{it}) + \beta' x_{it} + \delta_i + \delta_t + u_{it} \tag{3.9}$$

Results are shown in Table 3.3.[3] [4]

These results are not too convincing: the distance effect is only significant in the case of "Pulp, paper and paper products" and "Machinery and equipment, not elsewhere categorized". These significant coefficients are negative as expected; however, as motorways seem to have a significant effect only on two industries, these results do not really confirm my hypothesis about transportation-intensive and non-transportation-intensive industries (except for the fact that paper is indeed heavy and can be shipped in large quantities).



[3] I estimated these models with distance dummies as well as on a sample restricted to up to $10,000$ inhabitants. The first case did not lead to much different results: distance was still not significant. The restricted sample was too small to estimate in several cases.

[4] The same effects estimated using the first differencing method are shown in the Appendix, Table A.2.

**Table 3.3: The effect of motorways on TFP – separate FE estimations for each sector in the sample.**

| | FE: Log TFP | | |
|---|---|---|---|
| | Log distance | Obs. | $R^2$ |
| Textiles | -0.0851<br>(0.0674) | 3,015 | 0.092 |
| Wearing apparel; dressing and dyeing of fur | 0.0375<br>(0.0399) | 5,806 | 0.046 |
| Tanning and dressing of leather; luggage, handbags, [. . . ] | 0.0228<br>(0.0521) | 2,225 | 0.076 |
| Wood and of products of wood and cork, except furniture; [. . . ] | -0.0493<br>(0.0398) | 5,346 | 0.053 |
| Pulp, paper and paper products | -0.233**<br>(0.105) | 1,002 | 0.124 |
| Publishing, printing and reproduction of recorded media | 0.0563<br>(0.0449) | 3,627 | 0.036 |
| Rubber and plastic products | 0.00788<br>(0.0405) | 5,025 | 0.108 |
| Other non-metallic mineral products | -0.0455<br>(0.0486) | 3,551 | 0.111 |
| Basic metals | 0.0607<br>(0.0794) | 1,000 | 0.065 |
| Fabricated metal products, except machinery and equipment | 0.00875<br>(0.0235) | 12,113 | 0.067 |
| Machinery and equipment n.e.c. | -0.0464*<br>(0.0279) | 8,681 | 0.059 |
| Office machinery and computers | 0.0300<br>(0.101) | 410 | 0.043 |
| Electrical machinery and apparatus n.e.c. | 0.0872<br>(0.0583) | 2,729 | 0.168 |
| Radio, television and communication equipment and apparatus | 0.0951<br>(0.116) | 1,381 | 0.559 |
| Medical, precision and optical instruments, watches and clocks | -0.0263<br>(0.0485) | 2,101 | 0.077 |
| Motor vehicles, trailers and semi-trailers | -0.00646<br>(0.120) | 1,252 | 0.287 |
| Other transport equipment | 0.0609<br>(0.156) | 415 | 0.148 |
| Furniture; manufacturing n.e.c. | 0.0127<br>(0.0473) | 4,175 | 0.067 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Note*: All models include the following control variables: log number of employees, trader dummy, log distance to motorway × trader dummy, trader × log time to nearest border crossing; firm and year fixed effects are also included; Distance to motorway is calculated on the municipality level. See first difference estimations in the Appendix (Table A.2).

## 3.3 Difference in differences

Another estimation approach is using a difference in differences model. As I have mentioned in Chapter 1, this is also used to estimate the effects of motorways (Datta, 2012; Ghani, Goswami and Kerr, 2012; Márk, 2013), especially when there is a large change in the network and before–after comparisons are made.

The general idea behind the model is comparing two groups in the sample: those who got some treatment (in this case: better access to motorways) and those who did not. The results in the restricted sample fixed effects model (Table 3.2) has shown a somewhat significant increase in TFP for firms that were up to 20 kilometers away from the motorway. While this is not a highly robust result, I use this measure to classify firms: they are "treated" if they were farther than 20 kilometers away from the motorway but became closer than this limit. In order to have a control group that is relatively similar in market access, I used firms that were and stayed between 20-40 kilometers of the motorways (remember, this distance dummy was not significant in Table 3.2, i.e. these firms are just as affected as the ones even farther). Of course, this classification may seem arbitrary, and it would be better to have a clear-cut treatment/control setup. For example, when Card and Krueger (1994) compared employment in two American states, the treatment group was well defined: one state increased minimum wage and the other did not. However, both Datta (2012) and Márk (2013) used such arbitrary limits and found results that were both statistically and economically significant.

A naïve estimation of such an effect would simply compare the treated group before and after the treatment:

$$y_{it+k} = \beta_0 + \delta_t + u_i \tag{3.10}$$

where $i$ refers to the individual, $t$ is the date of the treatment, and $k$ is the required periods for the effect to fully develop; $\delta_t$ is a dummy that shows whether the individual is after treatment. This method is essentially just comparing the means of the treatment group before and after treatment.

I calculated these measures for all years in which motorways were built; I assumed $k = 1$, i.e. the effects (if they exist) are already there the next year. The results are shown in Table 3.4.

**Table 3.4: The treated group before and after treatment.**

| | Log(TFP) | | | | | |
|---|---|---|---|---|---|---|
| | 1994 | 1996 | 1997 | 1998 | 1999 | 2002 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | OLS | OLS | OLS | OLS | OLS | OLS |
| Post-treatment period | -0.230 | 0.0650 | 0.784 | 0.0715 | 0.357* | 0.166 |
| | (0.590) | (0.226) | (0.975) | (0.132) | (0.209) | (0.210) |
| Constant | -0.199 | 0.278* | -1.207 | 0.0369 | 0.290** | 0.418*** |
| | (0.259) | (0.162) | (0.918) | (0.0897) | (0.127) | (0.148) |
| Observations | 13 | 103 | 6 | 149 | 96 | 92 |
| R-squared | 0.015 | 0.001 | 0.139 | 0.002 | 0.030 | 0.007 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We can see that in most years the difference is not statistically significant except for 2002. However, this estimation omits a key variable: the underlying trend. If we do not control for this, it is possible that the intervention had no effect and the variable in question would have increased anyway. To include this trend – and then find the deviation from it – we need a control group that we also observe both before and after the treatment. Applying the method described by Wooldridge (2012, p.455) to the case of motorways, the effect that we are looking for is:

$$\hat{\delta} = (\overline{\log(TFP)}_{t+k,treat} - \overline{\log(TFP)}_{t+k,treat}) - (\overline{\log(TFP)}_{t,treat} - \overline{\log(TFP)}_{t,treat}) \tag{3.11}$$

which is the difference in differences (hence the name of the method): *if* we assume that even though the two groups could have started from different levels, they would have evolved similarly *and* we see that the difference between them has significantly changed after the treatment, *and* there was no other change that could have had a differential effect (e.g. changes in policy), *then* we know that this difference was caused by the treatment. The main strength of this model is this causal interpretation (when assumptions hold) and that it takes into consideration a change in the trend of the two groups. Before I discuss these assumptions and their validity in the present case, I show a version of the model that can be econometrically tested as well as its graphical explanation.

Wooldridge (2012, p.455) suggests estimating the model in the form

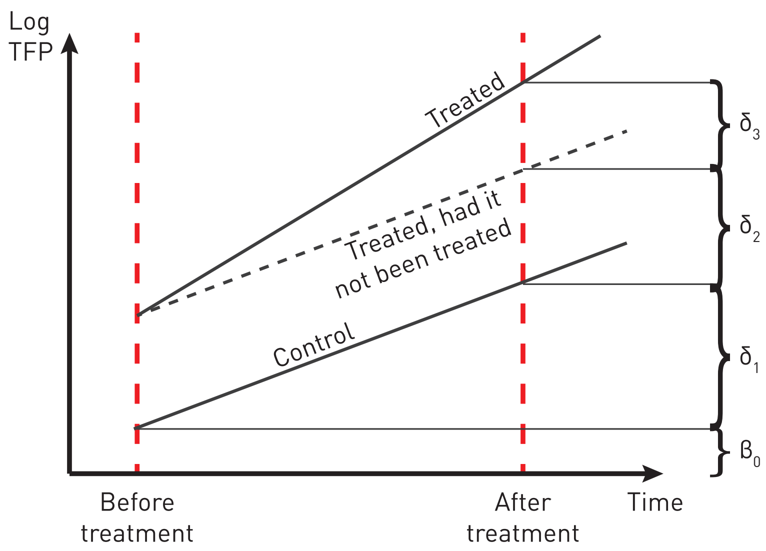
$$\log(TFP)_{it} = \beta_0 + \delta_1 aftertreatment_t + \delta_2 treated_i + \delta_3 (aftertreatment \times treated)_{it} + u_{it} \quad (3.12)$$

where *aftertreatment* and *treated* are dummies that indicate whether the observation is made in the post-treatment period and whether the observed firm was treated. Visually, these can be shown as in Figure 3.2: $\beta_0$ is a constant, $\delta_1$ is the common trend of the groups, $\delta_2$ is the initial difference between the two, and $\delta_3$ is the extra growth that happens to the treatment group by the post-treatment period. The common trend assumption is clearly important; if it does not hold, $\delta_3$ does not identify the effect correctly.

Since the data covers only a short time period, it is not really possible to check the common trend econometrically. Nevertheless, I checked visually on graphs whether the treated and control groups moved together. Such a graph for the treatment period of 1999 is shown in Figure 3.3. We can see that the values mostly move together – however, this cannot be taken a solid proof. It should also be noted that Hungary suffered a crisis around 1995 and a strict austerity package was introduced, which may have affected treated and control firms differently if they are not entirely similar in their characteristics. See, for example Table 3.5 for a comparison of control and treatment groups if the treatment year is 1999.

A general weakness of the difference-in-difference analysis here is that there are only few firms in the control group in each setup, which heavily hinders sector-level analysis or the re-

**Figure 3.2: The coefficients of a difference-in-differences estimation. Source: own work adapted from Márk (2013, p. 10)**




striction of the sample to settlements under $10,000$ inhabitants. Nevertheless, I ran these models; in many cases there were no treated firms at all in these restricted regressions, and even if there were a few, the effect was not significant. I do not show these results in tables.

Another factor to consider is that in the period that I analyze, there is no one single event that can be considered as a one time treatment; there are several years in which some motorways were built. Hence, I run this model on all such years. I also look at different lengths of $k$ (the time needed to achieve the effect after the intervention).

In Table 3.6 I present the results for $k = 1$, models of other values of $k$ are shown in the Appendix, Table A.3. The post-treatment dummy is significant in several of the cases, which suggests a general increase in TFP over time; the treatment dummy is significant in a few cases, which would suggest an initial difference between the two groups. The coefficient of interest, that of the interaction term, however, is not significant in any of the setups (for any $t$ and $k$). This implies, once again, that there is no significant effect of the motorways on productivity, even when looking at direct comparisons over time.

**Table 3.5: Comparison of the control and treatment group of 1999.**

| | *Control group in 1999* | | | | |
|---|---|---|---|---|---|
| | Mean | Count | Min | Max | SD |
| Empolyment | 93 | 1230 | 5 | 8427 | 347.30 |
| Distance to motorway | 29 | 1230 | 21 | 39 | 5.43 |
| Distance to Budapest | 88 | 1230 | 18 | 167 | 40.05 |
| Population of municipality | 23371 | 1230 | 498 | 62975 | 20453.95 |
| TFP | 0.32 | 1198 | -9.07 | 4.25 | 0.99 |

| | *Treatment group in 1999* | | | | |
|---|---|---|---|---|---|
| | Mean | Count | Min | Max | SD |
| Empolyment | 103 | 100 | 5 | 2003 | 283.77 |
| Distance to motorway | 20 | 100 | 1 | 50 | 14.26 |
| Distance to Budapest | 44 | 100 | 30 | 65 | 9.05 |
| Population of municipality | 17927 | 100 | 591 | 33592 | 15619.20 |
| TFP | 0.47 | 96 | -1.48 | 4.32 | 1.03 |

**Figure 3.3: Average log TFP of treated and control firms over time when the treatment year is 1999. Source: own calculations.**

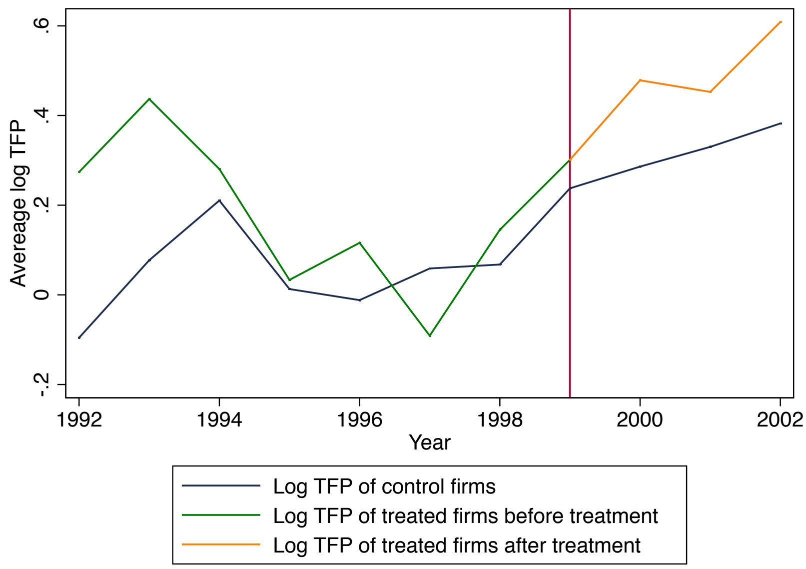


**Table 3.6: Difference in differences models by treatment year.**

| | Year of treatment ($t$). Dependent variable: log(TFP) | | | | | |
|---|---|---|---|---|---|---|
| | 1994 | 1996 | 1997 | 1998 | 1999 | 2002 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Post-treatment period | -0.0599 | 0.0844 | 0.0286 | 0.132*** | 0.149*** | 0.139*** |
| | (0.0577) | (0.0524) | (0.0512) | (0.0461) | (0.0495) | (0.0421) |
| Treatment | -0.494* | -0.0346 | -1.138 | -0.219** | 0.109 | 0.0357 |
| | (0.259) | (0.135) | (0.781) | (0.0872) | (0.112) | (0.120) |
| Post-treatment period × treatment | -0.168 | -0.0478 | 0.591 | -0.0834 | 0.170 | 0.0223 |
| | (0.581) | (0.191) | (0.882) | (0.120) | (0.165) | (0.166) |
| # treated | 7 | 55 | 3 | 75 | 50 | 48 |
| # control | 367 | 447 | 514 | 606 | 615 | 705 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Note:* The treatment group is the group of firms that were in a municipality that was between 20 and 40 km from the nearest motorway exit before the treatment year and became closer than 20 km in that year. Control group is the group of firms that were and stayed between 20 and 40 km until the post-treatment period. Pre-treatment period is $t-1$ in all models (i.e. one year before the firm became closer than 20 km); post-treatment period is $t+1$. Log number of employees and trader dummy are included in all estimations as well as a constant term.

# Chapter 4

# Start and cease of operation

In this chapter I analyze another effect that motorways may have on firms: first, I look at a municipal level to check whether there are more new firms born or closed, then I analyze individual firm decisions whether to cease operations by the next year. If market access is better, it can be more profitable to start a business, however, this market access also increases competition, because of which it is possible that more firms cease operations (or just that there are fewer new firms). It is not intuitively clear which effect should dominate; however, it would be preferred if motorways lead to more new firms and fewer closing ones.

## 4.1 Firm opening and closure on a municipality level

I analyze the number of firms that appear or disappear in the data each year. The data is not just a sample but includes all Hungarian firms that operate in these industries in this period and have more than 5 employees. I make the assumption that appearing in the data is in fact the birth of the firm, and when it disappears, it closes down. The data that is used in this section comes from the same sources as before, however, it is compiled in a different way: one observation is one municipality in a time period, where the explanatory variables are also on the municipality level.

Since the original data is on a firm-level, I can track the life cycle of each company individually. Among the more than $20,000$ firms, there are $3,050$ which only appear once in the data, i.e. they operated for not more than one year. I do not take these firms into consideration, because it is plausible that such firms were set up for specific projects and thus may systematically differ from "regular" firms, which I am interested in. Thus, I restrict the sample to firms that appear in at least three consecutive years thus ensuring that they operate for at least one full year. In this restricted group of firms there are $8,913$ new firms over the observed time period and $5,090$ that close down. I also exclude Budapest from the sample because it is an extreme outlier.

The general form of the model (based on Holl (2004*b*)) that I estimate is the following:

$$n_{it} = \beta_0 + \beta_1 f(dst_{it}) + \delta_t + \delta_t + u_{it} \tag{4.1}$$

where $n_{it}$ is the number of firms that start or close in settlement $i$ in year $t$. Once again, $f(\cdot)$ is either the logarithm or distance dummies. Further controls are the logarithm of time to the border and the number of inhabitants of the municipality; fixed effects are included for year and municipality. Holl (2004*a*) in a similar analysis includes wages, labor force qualification, several specialization and accessibility indices; these were not available for Hungary in the analyzed period and are thus omitted from my estimations.

Results are shown in Table 4.1. In columns (1)–(4) the dependent variable is the number of new firms that open in a given year in the municipality, in (5)–(8) it is the number of firms that close.

First, the effects are estimated using pooled OLS in (1) and (5): without town fixed effects but with year dummies. Most of the observations are 0 and since $\log(0)$ does not exist, we cannot take the logarithm of the dependent variable and thus cannot calculate elasticities. Linders and De Groot (2006) offer several solutions to such a problem when analyzing international trade:

**Table 4.1: Estimated models of the number of new firms and ones that cease operations.**

| | # new firms in municipality | | | | # closing firms in municipality | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) OLS | (2) FE | (3) Poisson | (4) Poisson | (5) OLS | (6) FE | (7) Poisson | (8) Poisson |
| Log distance to motorway | 0.0271** | 0.0252 | 0.0337 | | 0.0146** | 0.00153 | 0.0756 | |
| | (0.01306) | (0.0407) | (0.0605) | | (0.00739) | (0.0131) | (0.0957) | |
| *Distance to motorway...* | | | | | | | | |
| 0–29 km | | | | 0.00642 | | | | -0.329 |
| | | | | (0.167) | | | | (0.2128) |
| 20–39 km | | | | 0.0484 | | | | 0.0563 |
| | | | | (0.135) | | | | (0.1479) |
| Log time to border | -0.0926 | 0.636 | -1.845* | -1.844* | -0.0472*** | -0.407 | 0.285 | 0.0397 |
| | (0.01542) | (0.513) | (0.961) | (0.959) | (0.0085) | (0.292) | (1.827) | (1.5678) |
| Log number of inhabitants | 0.3585*** | -0.148** | -0.319 | -0.296 | 0.184*** | -0.103*** | -1.307** | -1.296 |
| | (0.0182) | (0.0700) | (0.526) | (0.523) | (0.0108) | (0.0346) | (0.645) | (0.6753) |
| Year fixed effects | Y | Y | Y | Y | Y | Y | Y | Y |
| Municipality fixed effects | N | Y | Y | Y | N | Y | Y | Y |
| Observations | 33,554 | 33,554 | 11,076 | 11,076 | 33,554 | 33,554 | 7,923 | 7,923 |
| $R^2$ | 0.1037 | 0.024 | | | 0.104 | 0.020 | | |
| Number of municipalities | 3,067 | 3,067 | 1,008 | 1,008 | | 3,067 | 721 | 721 |

Robust standard errors in parentheses*
*** p<0.01, ** p<0.05, * p<0.1

*Note:* Firms in the sample are those which operate for at least one whole year (i.e. we have observations on them for three years at least). Budapest is omitted from the estimation. Since we do not know whether firms that are in the database from the first year were born then or have already existed (and the case is similar with closure at the end of the data), birth is estimated between 1994–2002, firm closure between 1993–2001. In (4) and (8) the baseline distance group is 40+ km.
* Standard errors are not robust in (8) because Stata returned the following error trying to estimate the model: “variance matrix is nonsymmetric or highly singular”. Hence, normal standard errors are used.

imputing a small number, omitting zero observations, or using a two-step Heckman model.[1] Here I simply leave the number of firms as it is, even though it makes interpretation harder because this functional form would imply the same effect if the distance decreases from 200 to 100 or 2 to 1 kilometers, which is not very reasonable. For example, the coefficient (1) implies that in towns that are *ceteris paribus* 50% *farther from* the motorway, about $0.0271/100 \times 50 = 0.01355$ more firms are born each year – regardless of the original distance.

[1]This issue is very important in the trade literature since the model implies that logarithms of the variables must be used and it is not possible to ignore this as I did in this paper.

The coefficient of the distance is significantly different from zero both in (1) and (5) and both are positive, which means that cities closer to the motorway experience more firm openings and less closures. This implies that the expansion of competition prevents new entrants from entering the market, but it provides an edge for already existing firms so that they are less likely to close down. Across all time periods and municipalities, the average number of new firms in a year is 0.312 and the average distance of municipalities to the motorway is 90.88 kilometers; as stated above, decreasing this distance to its half, 45.44 kilometers would lead to the birth of 0.01355 more firms per settlement each year, which is about $0.01355 \times 3100 = 42$ new firms for the whole country in the whole year. This seems to be a rather negligible effect when compared to the significant costs of motorway investment.[2]

However, when municipality fixed effects are included in (2) and (6), even this small effect disappears; the estimated coefficient is smaller and insignificant in both estimations. This suggests that there are city-level fixed characteristics (e.g. location, city structure, etc.) that affect firm birth and closure but were not controlled for before.

Holl (2004*a*) uses a Poisson model to estimate a similar effect in Spain. According to Wooldridge (2012), this is a better method for the analysis of count data than linear regressions. Essentially, we estimate

$$y_{it} = \mathrm{e}^{\beta_0 + \beta' x_{it} + u_{it}} \tag{4.2}$$

using a maximum likelihood estimation. No matter whether the distance is included in the model as its logarithm or as distance-band dummies, it is insignificant for both firm birth and closure.

Holl (2004*a*) estimates separate models for each industry; this is not feasible in my dataset due to the low number of new and closing firms.[3] First, I tried running fixed-effects Poisson

[2] By 2006, the average distance from municipalities to the closest motorway ramp has actually reached 44.15 kilometers. As of 2013, it was 36 kilometers.

[3] More than 90 percent of the data is 0 firm births, about 9 percent is 1–5 new firms, and only the remaining 1 percent is more than 5. With firm closures, the distribution is even more skewed with more than 94 percent being zero and only 1 percent being 2 or more closures in a year in a municipality.

models for each industry, however, most of them did not converge. Then I simplified the estimation: the dependent variable became a binary one, it was 1 if there was at least new firm/firm that closed down in the given year. I tried estimating this model using a logit functional form with municipality fixed effect; convergence was achieved for most sectors, however, the effect of motorways was not significant in any of these models. Finally, I resorted to fixed-effect linear probability models, which yielded a few statistically significant estimates (at the 10 percent level), but these were not significant in the economic sense, because the coefficients were typically around $10^{-3}$ in magnitude, where the effect of a one percent decrease of distance to the motorway yields a probability change in the magnitude of $10^{-5}$. These results are shown in the Appendix in Table A.4.

## 4.2 The firm's decision to shut down

Since there is no way to know from hard data whether a firm wanted to open but did not, it is not possible to analyze such a decision directly, only on some level of spatial aggregation – for example, municipalities, as in the previous section. However, the end of the life of a firm – whether it closes or not – is a decision that we can observe directly. I return to the micro-level database in this section and show a model for the decision to shut down.

A firm can only operate in the long term if its profit is nonnegative, i.e. its income is higher than the costs it incurs. It can also be assumed that there is a positive relationship between the firm's profit and its survival until next year.

So profits are

$$\Pi = pq - c(q,X) \tag{4.3}$$

where $X$ is a vector of other factors, for example, in my database, market access, productivity of the firm, whether it trades or not, population of the municipality (since it may affect wages) are such effects. Then, since $\Pr(close) = f(\Pi)$,

$$\Pr(close) = f(X) \tag{4.4}$$

if we hold $q$ constant, or with a convenient linear functional form, we can divide both sides by $q$ and just look at $\frac{\Pi}{q}$.

To assess whether there is an effect of motorways on the firm's decision to shut down, I estimate the following model:

$$\Pr(close)_{it} = g(f(dst)_{it}, X, \delta_t, \delta_i) \tag{4.5}$$

where $f(\cdot)$ is the usual logarithmic or categorical distance, $X$ contains TFP, number of employees, trader status, and log population of the municipality; $\delta_t$ and $\delta_i$ are time and firm fixed effects. The actual dependent variable is a dummy: it's 1 if the firm does not appear in the database the next year and 0 if it does. Observations from 2003 are not included as we do not know whether firms stay active after our analysis period. Ideally, such a model would be estimated using a conditional logit model (conditional on the firm), however, probably due to the relatively few observations for each firm and their large numbers, such nonlinear estimations did not converge. Thus, I resorted to using the following linear model and its variations:

$$\Pr(close)_{it} = \beta_0 + \beta_1 f(dst)_{it} + \beta' X + \delta_t + \delta_i + u_{it} \tag{4.6}$$

Results are shown in Table 4.2.

When fixed effects are omitted in the OLS estimations of Column (1) and (2), the effects are significant and negative (as before in the Poisson municipality level estimations): more competition resulting form a larger market may be connected with closing the firm. It is not a result of this estimation, but the data tells us that across all years there is a 10 percent chance of a firm not

**Table 4.2: Estimated models of the probability that the firm ceases operations by next year.**

| | Firm ceases operation by next year (1 = yes, 0 = no | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | OLS | OLS | FE | FE | FE | FE |
| Log distance to motorway | -0.00425*** | | 0.000526 | 0.000051 | | |
| | (0.000979) | | (0.00524) | (0.00526) | | |
| *Distance to motorway...* | | | | | | |
| 0–19 km | | 0.0127*** | | | -0.01000 | -0.00776 |
| | | (0.00292) | | | (0.0135) | (0.0136) |
| 20–39 km | | 0.0101*** | | | 0.0111 | 0.0131 |
| | | (0.00358) | | | (0.0112) | (0.0112) |
| Log TFP | -0.0454*** | | -0.0347*** | -0.0359*** | -0.0347*** | -0.0359*** |
| | (0.00163) | | (0.00281) | (0.00292) | (0.00281) | (0.00292) |
| Log number of employees | -0.00985*** | | -0.0706*** | -0.0688*** | -0.0706*** | -0.0688*** |
| | (0.00109) | | (0.00355) | (0.00354) | (0.00355) | (0.00354) |
| Trader | -0.0242*** | | -0.0133*** | -0.0122*** | -0.0134*** | -0.0123*** |
| | (0.00281) | | (0.00409) | (0.00410) | (0.00409) | (0.00410) |
| Log population of municipality | 0.00426*** | | 0.00323 | 0.0175 | 0.00479 | 0.0193 |
| | (0.000780) | | (0.0415) | (0.0414) | (0.0414) | (0.0414) |
| Year fixed effect | N | N | Y | N | Y | N |
| Year $\times$ sector fixed effect | N | N | N | Y | N | Y |
| Constant | Y | Y | Y | Y | Y | Y |
| Joint test of distance dummies ($F(10,11615)$) | | | | | 1.1 | 1.07 |
| p-value | | | | | 0.3544 | 0.3837 |
| Observations | 57,776 | 57,776 | 57,776 | 57,776 | 57,776 | 57,776 |
| $R^2$ (within) | 0.036 | 0.036 | 0.103 | 0.109 | 0.103 | 0.109 |
| Firms in sample | | | 11,616 | 11,616 | 11,616 | 11,616 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Note:* I tried running fixed-effect logit models (`clogit` and `xtlogit` commands in Stata), however, these did not converge – probably due to the large number of groups in the estimation. In (2), (6), and (7) the baseline distance group is 0–9 km. Distance to motorway is calculated on the municipality level.

appearing in the database next year (and thus being categorized as closed). The average distance of firms (again, taking all years into consideration) is 46 kilometers. This suggests that a firm at about 100 kilometers from the motorway is about 0.4 percent less likely to close compared to the firm at 46 kilometers – this effect is rather small. Adding fixed effects (for each firm and for year or year-sector interaction) to filter out these effects yields even smaller and insignificant coefficients regardless of the functional form of the distance.

However, in this case it is interesting to look at the other coefficients as well: "better functioning firms" (those with higher TFP) are less likely to close; and this is true for larger firms as well as those who trade – which is an intuitively sensible result.

# Chapter 5

# Conclusion

In this thesis I analyzed the effect of motorways on firm performance: productivity, the number new firms and firms that close in a municipality, and the individual firm's decision whether to cease operations. I used several methods and looked at various functional forms to estimate the effects, however, except for a fixed effect model restricted to small settlements, no significant effects emerged in proper models. When the effects were significant, they were due to omitted variables. From these results, if there is any policy recommendation to make, it is that motorways should not be built just because of their expected benefit to be a significant increase in manufacturing productivity – this benefit likely does not exist, or is relatively modest and only applies to small rural settlements.

This, even if this result is universally correct, not just for the sample in question, does not mean that motorways have no benefits. First, while it is plausible that manufacturing firms benefit significantly from better transportation, this can also be thought about other sectors as well, for example, tourism. My data set only contained manufacturing firms of different kinds (about $20,000$ of the $600,000$ firms that operate in Hungary), thus I cannot test the performance of other firms and sectors. Since 2003, the total length of motorways have doubled. Effects may arise now if a full network of motorways has greater benefits then sorter segments that do not reach the border. It is also possible that motorways now reach parts of the country that benefit more from better access to markets (for example because market access was so abysmal before). It has been shown that in developing countries the construction of rural roads has immense benefits (Jacoby, 2000; Jacoby and Minten, 2009); while it is true that every Hungarian municipality can be reached on asphalt roads, there are lots of roads whose quality could be vastly improved, for example by adding new lanes or bypasses around villages.

Other types of benefits can arise on roads as well. Most importantly, time savings and the value of avoided accidents is large and significant (European Court of Auditors, 2013). Uses other than freight transportation can beneficial as well: for example, public transportation that uses this road or people who can now go to work because travel has became faster and safer.

I have not been able to show significant effects of motorways on productivity in Hungary. However, further research on panel data with more time periods and a wider range of firms may be able to demonstrate such effects.

# Bibliography

**bama.hu.** 2010. “Avatás, átadás, ünnep: az M6/M60 napja [Opening ceremony and celebration: the day of Motorway M6/M60].” Available at: `https://web.archive.org/web/20130922114310/http://www.bama.hu/baranya/kozelet/avatas-atadas-unnep-az-m6m60-napja-videok-296236` Last accessed: June 1, 2014.

**Békés, Gábor, and Péter Harasztosi.** 2013. “Agglomeration premium and trading activity of firms.” *Regional Science and Urban Economics*, 43(1): 51–64.

**Békés, Gabor, Balázs Muraközy, and Péter Harasztosi.** 2011. “Firms and products in international trade: Evidence from Hungary.” *Economic Systems*, 35(1): 4–24.

**Boarnet, Marlon G.** 1995. “Highways and economic productivity: interpreting recent evidence.” UCTC Working Paper No. 291.

**Bontemps, Christophe, Michel Simioni, and Yves Surry.** 2008. “Semiparametric hedonic price models: assessing the effects of agricultural nonpoint source pollution.” *Journal of Applied Econometrics*, 23(6): 825–842.

**Brakman, Steven, Harry Garretsen, and Charles van Marrewijk.** 2009. *The new introduction to geographical economics.* Cambridge University Press.

**Bronzini, Raffaello, and Paolo Piselli.** 2009. "Determinants of long-run regional productivity with geographical spillovers: the role of R&D, human capital and public infrastructure." *Regional Science and Urban Economics*, 39(2): 187–199.

**Brülhart, Marius, and Federica Sbergami.** 2009. "Agglomeration and growth: Cross-country evidence." *Journal of Urban Economics*, 65(1): 48–63.

**Brülhart, Marius, Matthieu Crozet, and Pamina Koenig.** 2004. "Enlargement and the EU periphery: the impact of changing market potential." *The World Economy*, 27(6): 853–875.

**Card, David, and Alan B. Krueger.** 1994. "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." *The American Economic Review*, 84(4): pp. 772–793.

**Chandra, Amitabh, and Eric Thompson.** 2000. "Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system." *Regional Science and Urban Economics*, 30(4): 457–490.

**Ciccone, Antonio, and Robert E. Hall.** 1996. "Productivity and the Density of Economic Activity." *The American Economic Review*, 86(1): pp. 54–70.

**Datta, Saugato.** 2012. "The impact of improved highways on Indian firms." *Journal of Development Economics*, 99(1): 46–57.

**Davis, Donald R, and David E Weinstein.** 2001. "Market size, linkages, and productivity: a study of Japanese regions." NBER Working Paper 8518.

**European Court of Auditors.** 2013. "Are EU Cohesion Policy funds well spent on roads?" special report no 5/2013.

**Fujita, Masahisa, Paul R. Krugman, and Anthony J. Venables.** 1999. *The Spatial Economy.* MIT Press.

**Ghani, Ejaz, Grover Goswami, and William R. Kerr.** 2012. "Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing." Harvard Business School Harvard Business School Working Papers 13-040.

**Holl, Adelheid.** 2004*a*. "Manufacturing location and impacts of road transport infrastructure: empirical evidence from Spain." *Regional Science and Urban Economics*, 34(3): 341–363.

**Holl, Adelheid.** 2004*b*. "Transport Infrastructure, Agglomeration Economies, and Firm Birth: Empirical Evidence from Portugal*." *Journal of Regional Science*, 44(4): 693–712.

**Holl, Adelheid.** 2007. "Twenty years of accessibility improvements. The case of the Spanish motorway building programme." *Journal of Transport Geography*, 15(4): 286–297.

**Holl, Adelheid.** 2012. "Market potential and firm-level productivity in Spain." *Journal of Economic Geography*, 12(6): 1191–1215.

**Holl, Adelheid.** 2014. "Highways and productivity in urban and rural locations." Working paper.

**Hornung, Erik.** 2012. "Railroads and micro-regional growth in prussia." Ifo Working Paper 127.

**Jacoby, Hanan G.** 2000. "Access to markets and the benefits of rural roads." *The Economic Journal*, 110(465): 713–737.

**Jacoby, Hanan G, and Bart Minten.** 2009. "On measuring the benefits of lower transport costs." *Journal of Development Economics*, 89(1): 28–38.

**Koster, Hans RA, and Jan Rouwendal.** 2012. "The impact of mixed land use on residential property values." *Journal of Regional Science*, 52(5): 733–761.

**Krugman, Paul R.** 1991. *Geography and trade.* MIT press.

**Lileeva, Alla, and Daniel Trefler.** 2010. "Improved access to foreign markets raises plant-level productivity... for some plants." *The Quarterly Journal of Economics*, 125(3): 1051–1099.

**Linders, Gert-Jan M, and Henri LF De Groot.** 2006. "Estimation of the gravity equation in the presence of zero flows." Tinbergen Institute Discussion Paper TI 2006-072/3.

**Lokshin, Michael.** 2006. "Difference-based semiparametric estimation of partial linear regression models." *The Stata Journal*, 6(3): 377–383.

**Madura, Máté.** 2013. "Motorway segments in Hungary and their date of opening." Unpublished.

**Márk, Lili.** 2013. "The Effect of Highways on Nearby Residential Property Prices in Hungary." Master's thesis. Central European University, Budapest.

**Martín-Barroso, David, Juan A Núñez, and Francisco J Velázquez.** 2013. "The Effect On Firms? Productivity Of Accessibility. The Spanish Manufacturing Sector." European Regional Science Association.

**Melo, Patricia C., Daniel J. Graham, and Ruben Brage-Ardao.** 2013. "The productivity of transport infrastructure investment: A meta-analysis of empirical evidence." *Regional Science and Urban Economics*, 43(5): 695 – 706.

**Németh, Nándor.** 2005. "Az autópálya-hálózat térszerkezet alakító hatásai – Magyarország esete [The effect of the motorway network on spatial structures – the case of Hungary]." In *Munkapiac és rergionalitás Magyarországon.* , ed. K. Fazekas, 139–179. MTA.

**Ohnsorge-Szabó, László.** 2006. "Közlekedési infrastruktúra és jólét Kelet-Magyarországon [Transport infrastructure and welfare in eastern Hungary]." *Statisztikai Szemle*, 84(3): 249–270.

**Olley, G. Steven, and Ariel Pakes.** 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64(6): pp. 1263–1297.

**osm.org.** 2014. "OpenStreetMap." Available at `http://openstreetmap.org`. Last accessed: May 28, 2014.

**Österreichisches Staatsarchiv.** 2014. “The Second Military Survey.” `http://mapire.eu/en/secondsurvey/`, Accessed: 2014-05-20.

**Pradhan, Rudra P, and Tapan P Bagchi.** 2013. “Effect of transportation infrastructure on economic growth in India: The VECM approach.” *Research in Transportation Economics*, 38(1): 139–148.

**Rephann, Terance, and Andrew Isserman.** 1994. “New highways as economic development tools: An evaluation using quasi-experimental matching methods.” *Regional Science and Urban Economics*, 24(6): 723–751.

**Wooldridge, Jeffrey M.** 2012. *Introductory Econometrics.* . 5 ed., South-Western.

# Appendix A

# Tables

**Table A.1: Nth difference models of relationship between distance to motorways and productivity.**

| | $\Delta_i$ Log TFP | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | 1st diff ($i = 1$) | 1st diff ($i = 1$) | 2nd diff ($i = 2$) | 3rd diff ($i = 3$) |
| $\Delta_i$ Log distance to motorway | -0.00152 | -0.00446 | -0.00584 | -0.0125 |
| | (0.0114) | (0.0114) | (0.0141) | (0.0166) |
| Year fixed effect | Y | N | N | N |
| Year $\times$ sector fixed effect | N | Y | Y | Y |
| Other controls | Y | Y | Y | Y |
| Constant | Y | Y | Y | Y |
| Observations | 50,555 | 50,555 | 40,073 | 31,639 |
| R-squared | 0.011 | 0.025 | 0.039 | 0.053 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Note*: $\Delta_i x \equiv x_t - x_{t-i}$. "Other controls" are: log number of employees, trader dummy, log distance to motorway $\times$ trader dummy, trader $\times$ log time to nearest border crossing. Distance to motorway is calculated on the municipality level.

**Table A.2: The effect of motorways on TFP – separate FD estimations for each sector in the sample.**

| | FD: Δ log TFP | | |
|---|---|---|---|
| | Δ log distance | Obs. | $R^2$ |
| Textiles | -0.0642<br>(0.0616) | 2,317 | 0.013 |
| Wearing apparel; dressing and dyeing of fur | 0.0196<br>(0.0389) | 4,652 | 0.019 |
| Tanning and dressing of leather; luggage, handbags, [. . . ] | 0.0761<br>(0.0611) | 1,773 | 0.041 |
| Wood and of products of wood and cork, except furniture; [. . . ] | -0.0817**<br>(0.0415) | 4,086 | 0.014 |
| Pulp, paper and paper products | -0.247**<br>(0.108) | 807 | 0.036 |
| Publishing, printing and reproduction of recorded media | 0.00605<br>(0.0409) | 2,817 | 0.039 |
| Rubber and plastic products | 0.0487<br>(0.0398) | 3,985 | 0.033 |
| Other non-metallic mineral products | -0.0236<br>(0.0570) | 2,816 | 0.025 |
| Basic metals | 0.0149<br>(0.106) | 793 | 0.037 |
| Fabricated metal products, except machinery and equipment | 0.0253<br>(0.0253) | 9,619 | 0.027 |
| Machinery and equipment n.e.c. | -0.0461<br>(0.0305) | 6,968 | 0.020 |
| Office machinery and computers | 0.166**<br>(0.0767) | 324 | 0.041 |
| Electrical machinery and apparatus n.e.c. | 0.00965<br>(0.0402) | 2,221 | 0.047 |
| Radio, television and communication equipment and apparatus | 0.0856<br>(0.0880) | 1,065 | 0.072 |
| Medical, precision and optical instruments, watches and clocks | -0.0294<br>(0.0499) | 1,698 | 0.034 |
| Motor vehicles, trailers and semi-trailers | -0.00772<br>(0.103) | 1,023 | 0.038 |
| Other transport equipment | 0.146<br>(0.161) | 318 | 0.032 |
| Furniture; manufacturing n.e.c. | -0.0583<br>(0.0502) | 3,273 | 0.012 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Note*: All models include the following control variables: log number of employees, trader dummy, log distance to motorway × trader dummy, trader × log time to nearest border crossing; year dummies are also included. Distance to motorway is calculated on the municipality level.

**Table A.3: Difference in differences models by treatment year.**

| | | Year of treatment ($t$). Dependent variable: log(TFP) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Post-treatment period | | 1994<br>(1) | 1996<br>(2) | 1997<br>(3) | 1998<br>(4) | 1999<br>(5) | 2002<br>(6) | 2003<br>(7) |
| | Post-treatment period | 0.00179 | 0.0448 | 0.0438 | 0.0179 | 0.125*** | 0.116*** | 0.0123 |
| | | (0.0567) | (0.0545) | (0.0493) | (0.0466) | (0.0435) | (0.0426) | (0.0410) |
| | Treatment | -0.833** | -0.00260 | -1.120 | -0.189** | 0.124 | 0.0341 | -0.159** |
| | | (0.341) | (0.129) | (0.781) | (0.0871) | (0.102) | (0.114) | (0.0642) |
| $t$ | Post-treatment period × treatment | -0.148 | 0.0443 | 0.524 | -0.0936 | -0.0571 | -0.0127 | 0.0577 |
| | | (0.711) | (0.173) | (0.946) | (0.119) | (0.150) | (0.159) | (0.0868) |
| | # treated | 10 | 58 | 3 | 82 | 65 | 51 | 256 |
| | # control | 417 | 492 | 550 | 649 | 778 | 761 | 832 |
| | Post-treatment period | -0.0263 | 0.0453 | 0.156*** | 0.192*** | 0.152*** | | |
| | | (0.0609) | (0.0548) | (0.0534) | (0.0539) | (0.0484) | | |
| | Treatment | -0.390 | -0.0711 | -1.220 | -0.220** | 0.217** | | |
| | | (0.272) | (0.141) | (0.791) | (0.0943) | (0.107) | | |
| $t+2$ | Post-treatment period × treatment | -0.0136 | -0.0426 | 0.682 | -0.123 | -0.0523 | | |
| | | (0.377) | (0.183) | (0.887) | (0.131) | (0.166) | | |
| | # treated | 6 | 52 | 3 | 64 | 45 | | |
| | # control | 318 | 417 | 482 | 488 | 587 | | |
| | Post-treatment period | 0.00520 | 0.182*** | 0.181*** | 0.173*** | 0.223*** | | |
| | | (0.0597) | (0.0549) | (0.0594) | (0.0548) | (0.0507) | | |
| | Treatment | -0.429 | -0.0737 | -1.262 | -0.242** | 0.200* | | |
| | | (0.278) | (0.151) | (0.795) | (0.0955) | (0.108) | | |
| $t+3$ | Post-treatment period × treatment | 0.142 | 0.0656 | 0.719 | -0.135 | 0.0909 | | |
| | | (0.297) | (0.202) | (0.845) | (0.136) | (0.171) | | |
| | # treated | 6 | 47 | 3 | 64 | 44 | | |
| | # control | 292 | 391 | 396 | 462 | 547 | | |
| | Post-treatment period | 0.0170 | 0.189*** | 0.170*** | 0.267*** | 0.271*** | | |
| | | (0.0617) | (0.0638) | (0.0599) | (0.0584) | (0.0506) | | |
| | Treatment | -0.427 | -0.0942 | -1.710 | -0.235** | 0.193* | | |
| | | (0.272) | (0.161) | (1.085) | (0.0959) | (0.109) | | |
| $t+4$ | Post-treatment period × treatment | 0.136 | 0.146 | 1.088 | -0.174 | 0.0504 | | |
| | | (0.295) | (0.216) | (1.124) | (0.152) | (0.180) | | |
| | # treated | 6 | 44 | 2 | 62 | 44 | | |
| | # control | 276 | 329 | 379 | 433 | 507 | | |
| | Post-treatment period | 0.119* | 0.202*** | 0.266*** | 0.301*** | | | |
| | | (0.0619) | (0.0634) | (0.0634) | (0.0575) | | | |
| | Treatment | -0.425 | -0.142 | -1.810* | -0.239** | | | |
| | | (0.281) | (0.165) | (1.087) | (0.100) | | | |
| $t+5$ | Post-treatment period × treatment | -0.0256 | 0.144 | 1.228 | -0.123 | | | |
| | | (0.441) | (0.208) | (1.089) | (0.145) | | | |
| | # treated | 6 | 41 | 2 | 58 | | | |
| | # control | 258 | 314 | 356 | 402 | | | |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Note:* The treatment group is the group of firms that were in a municipality that was between 20 and 40 km from the nearest motorway exit before the treatment year and became closer than 20 km in that year. Control group is the group of firms that were and stayed between 20 and 40 km until the post-treatment period. Pre-treatment period is $t-1$ in all models (i.e. one year before the firm became closer than 20 km). Log number of employees and trader dummy are included in all estimations as well as a constant term.

**Table A.4: Separate fixed effect linear estimations of the probability that there is at least one firm opening/closing in the municipality in the given year for each sector.**

| | Coefficient of distance to motorway... | | | | |
|---|---|---|---|---|---|
| Sector: | 0–19 km | | 20–39 km | | R-squared |
| *a) There is at least one firm that opens in the municipality (1 = yes, 0 = no)* | | | | | |
| 17 | -0.00356* | (0.00187) | 0.000112 | (0.00134) | 0.011 |
| 18 | -0.00481* | (0.00252) | 0.000889 | (0.00132) | 0.013 |
| 19 | -0.00206 | (0.00155) | -0.000116 | (0.000816) | 0.013 |
| 20 | -0.00399 | (0.00253) | 0.00132 | (0.00151) | 0.016 |
| 21 | -0.00154 | (0.00129) | 0.000375 | (0.000544) | 0.012 |
| 22 | -0.00162 | (0.00204) | 0.00237** | (0.00115) | 0.011 |
| 23 | 5.05e-05 | (3.24e-05) | 7.55e-05*** | (2.58e-05) | 0.036 |
| 25 | -0.00294 | (0.00285) | 0.00168 | (0.00140) | 0.014 |
| 26 | -0.00186 | (0.00211) | 0.000817 | (0.00134) | 0.011 |
| 27 | -0.00221 | (0.00141) | 0.00152** | (0.000705) | 0.015 |
| 28 | -0.00548 | (0.00398) | 0.00574** | (0.00233) | 0.026 |
| 29 | -0.00394 | (0.00308) | 0.00241 | (0.00155) | 0.018 |
| 30 | 0.000426 | (0.000616) | 0.000263 | (0.000358) | 0.034 |
| 31 | -0.00277 | (0.00179) | 0.00194** | (0.000868) | 0.011 |
| 32 | -0.000360 | (0.00118) | 0.00237*** | (0.000713) | 0.012 |
| 33 | -0.00229 | (0.00179) | 0.00182** | (0.000770) | 0.011 |
| 34 | 0.000182 | (0.00111) | 0.00125** | (0.000608) | 0.013 |
| 35 | -0.000637 | (0.000705) | 0.000408 | (0.000331) | 0.018 |
| 36 | -0.00332 | (0.00216) | 0.000823 | (0.00132) | 0.013 |
| 37 | -0.000456 | (0.000637) | 0.000654 | (0.000397) | 0.026 |
| *b) There is at least one firm that closes the municipality (1 = yes, 0 = no)* | | | | | |
| 17 | 0.0505*** | (0.0160) | -0.0233* | (0.0126) | 0.419 |
| 18 | -0.00392* | (0.00212) | 0.000736 | (0.00110) | 0.011 |
| 19 | -0.00319* | (0.00174) | -0.000694 | (0.000740) | 0.005 |
| 20 | -0.00299 | (0.00191) | 0.000291 | (0.00121) | 0.009 |
| 21 | -0.000832 | (0.000561) | 6.65e-05 | (0.000334) | 0.006 |
| 22 | -0.00291* | (0.00174) | 0.00195*** | (0.000724) | 0.006 |
| 23 | 0.000121* | (6.25e-05) | 9.36e-05** | (4.44e-05) | 0.001 |
| 25 | -0.00120 | (0.00145) | 0.00102 | (0.000841) | 0.007 |
| 26 | -0.00211 | (0.00133) | 6.08e-05 | (0.000829) | 0.007 |
| 27 | -0.00127 | (0.00102) | 0.000543 | (0.000407) | 0.004 |
| 28 | -0.00372 | (0.00253) | 0.00254* | (0.00131) | 0.014 |
| 29 | -0.00470* | (0.00251) | 0.000978 | (0.00105) | 0.012 |
| 30 | -0.000401 | (0.000449) | 0.000117 | (0.000217) | 0.004 |
| 31 | -0.00168 | (0.00113) | 0.000935** | (0.000472) | 0.006 |
| 32 | -3.15e-05 | (0.000561) | 0.000689 | (0.000455) | 0.005 |
| 33 | -0.00127 | (0.00125) | 0.000990 | (0.000660) | 0.004 |
| 34 | 0.000764* | (0.000427) | 0.000656 | (0.000486) | 0.005 |
| 35 | -2.10e-05 | (0.000423) | 0.000227 | (0.000315) | 0.003 |
| 36 | -0.00189 | (0.00135) | 0.000209 | (0.00102) | 0.007 |
| 37 | 7.68e-05 | (0.000222) | 0.000240 | (0.000181) | 0.004 |

Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

*Note:* All estimations are linear probability estimations run on 68,839 observations of 3,165 municipalities; they include year and municipality fixed effects as well as the population of the municipality as an additional control variable. Firms are labeled accoding to NACE 1.1 classification: 17: Manufacture of textiles; 18: Manufacture of wearing apparel; dressing and dyeing of fur; 19: Tanning and dressing of leather; manufacture of luggage, handbags, saddlery, harness and footwear; 20: Manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials; 21: Manufacture of pulp, paper and paper products; 22: Publishing, printing and reproduction of recorded media; 23: Manufacture of coke, refined petroleum products and nuclear fuel; 24: Manufacture of chemicals and chemical products; 25: Manufacture of rubber and plastic products; 26: Manufacture of other non-metallic mineral products; 27: Manufacture of basic metals; 28: Manufacture of fabricated metal products, except machinery and equipment; 29: Manufacture of machinery and equipment n.e.c.; 30: Manufacture of office machinery and computers; 31: Manufacture of electrical machinery and apparatus n.e.c.; 32: Manufacture of radio, television and communication equipment and apparatus; 33: Manufacture of medical, precision and optical instruments, watches and clocks; 34: Manufacture of motor vehicles, trailers and semi-trailers; 35: Manufacture of other transport equipment; 36: Manufacture of furniture; manufacturing n.e.c.; 37: Recycling.

# Appendix B

# Figures

**Figure B.1: Travel time to Budapest in 1992 and 2012. Source: own work based on OpenStreetMap data.**