TI 2017-109/VIII
Tinbergen Institute Discussion Paper

The impact of highways on population redistribution: The role of land development restrictions

Or Levkovich[1]
Jan Rouwendal[1,2]
Jos van Ommeren[1 2]

[1] Department of Spatial Economics, VU University, The Netherlands
[2] Tinbergen Institute, The Netherlands

The impact of highways on population redistribution: The role of land development restrictions

Or Levkovich[a,*], Jan Rouwendal[a,b], Jos van Ommeren[a,b]

[a] *Department of Spatial Economics, VU University, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.*
[b] *Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands.*

Abstract

We study the role of land development restrictions for the effects of highway expansion on the spatial distribution of population. We demonstrate that these restrictions strongly interfered with the effects of highways in the Netherlands. Introducing an IV approach to address endogenous interaction variables, our findings show that new highways accelerated population growth in peripheral areas, but had no such effect in central cities and suburban municipalities. We find that due to development restrictions near larger cities, the highway expansion caused a 'leapfrog' pattern, in which suburban growth skipped development-restricted areas and expanded into farther located peripheral areas.

Keywords: highways, development restrictions, population redistribution, suburbanization, instrumental variables, endogenous interaction variables

---

1. Introduction

New highways transform the spatial structure of cities and regions by reducing the costs of commuting to employment centers and improving accessibility in peripheral areas. There is substantial evidence that highways induce suburbanization, reduce population densities in central cities and increase population levels and economic performance of peripheral areas (Baum-Snow, 2007a,b; Chandra and Thompson, 2000; Duranton and Turner, 2012; Garcia-López et al., 2015). However, these predictions may change considerably where urban sprawl is bounded by land development restrictions, prevalent in many cities around the world.

Land development restrictions are often introduced to mitigate urban sprawl. The green belt surrounding London is a well-known example, but similar development restrictions exist in the surroundings of many other cities.[1] Our focus is on the Netherlands, where the protection of the inner part of the Randstad, a metropolitan area which contains the four largest cities, against urbanization (through the so-called 'Green Heart') is an important aspect of the planning

---

*Corresponding author.
Email addresses: o.d.Levkovich@vu.nl (Or Levkovich), j.rouwendal@vu.nl (Jan Rouwendal), jos.van.ommeren@vu.nl (Jos van Ommeren)

[1] Notable examples include the Ontario and Ottawa green belts, the São Paolo Biosphere reserve and the Seoul green belt (Jun and Hur, 2001).

system. In addition, buffer zones contiguous to these cities were introduced to ensure proximity to open space and prevent the merging of urbanized areas.

Not much is known about the effect of urban planning measures on the spatial distribution of the population. There is abundant evidence that land development restrictions increase house prices (Glaeser et al., 2006; Kok et al., 2014), and can impose limitations on employment growth (Hsieh and Moretti, 2017). This implies that restrictions are relevant and that restrictions on land use in the vicinity of large cities induce people to reside further away.

Following the work of Baum-Snow (2007a) it is well known that highway construction was responsible for a substantial share of the suburbanization of US cities. Most of this suburbanization took place through gradual expansion of the urbanized area and increasing densification within the municipal boundaries (Burchfield et al., 2006). In this paper, we examine how the growth of the highway network in the Netherlands interacted with land development restrictions to transform the spatial distribution of population. More specifically, we address the question whether land development restrictions resulted in different population growth rates than could otherwise be expected on the basis of the expanding highway network.

To determine the causal effect of highway network developments on suburbanization, Baum-Snow used an instrument based on historical plans for network development. Later studies have used similar historical instruments to study, for instance, the effect of highway networks on Chinese cities (Baum-Snow et al., 2017), on land conversion in Spain and on urban structure in Barcelona (Garcia-López et al., 2014, 2015), on innovation in US regions (Agrawal et al., 2017) and on employment levels in Italian cities (Percoco, 2016).

We focus on a large scale expansion of the Dutch highway network in the 1960's, and study its impact on municipal population growth during the decade that followed. We account for endogeneity issues by employing an instrumental variable approach using the 1821 road network. We find that highways have contributed to population growth in peripheral municipalities, but not in suburbs or central cities. Development restrictions resulted in a 'leapfrog' of suburbanization over areas in which development is restricted, population growth primarily occurred in peripheral towns. We also find that the effect of highways on population growth is temporal, as highways seem to have limited effect on population growth more than ten years after the major expansion is completed.

In our empirical strategy, we pay special attention to the treatment of endogenous interaction variables. We use an innovative econometric approach in which we use a single instrument and estimate a single first-stage regression to address multiple interaction terms between an endogenous and exogenous variables. Initially suggested by Balli and Sorensen (2013), we further develop the single first-stage approach and formally show that when the exogenous interaction variable is a categorical variable (which splits the sample into exclusive subsamples), it produces more efficient results than the commonly-applied multiple first-stage approach, under certain testable conditions. Furthermore, we will point out that this approach is biased, in our context at least, due to weak instruments.

The paper is organized as follows. Section 2 provides a theoretical background. Section 3 discusses important aspects of Dutch land use planning and road network development. Section 4 describes the data. Section 5 describes our methodological approach. Section 6 includes the estimation results. Section 7 includes additional sensitivity analysis and examines long-term effects. Section 8 concludes.

2. Theoretical background

Following the monocentric model, the development of road infrastructure has a positive effect on city size and suburbanization (Alonso, 1964; Muth, 1969; Mills, 1967). In the closed city version of the model, the rent gradient flattens with a decrease in transportation costs, land in the center becomes less expensive and the city expands through larger average lot sizes. The open city version of the model also predicts a flatter rent gradient, and a larger population. The Roback (1982) model suggests a positive effect of lower local transportation costs on employment on top of that. Glaeser and Kahn (2004) argued that reduction in commuting costs, which accompanied the widespread use of car transport, is the most important driver of suburbanization. This argument was also affirmed by others, notably by Burchfield et al. (2006).

Anas and Moses (1979) and Baum-Snow (2007b) considered an extension of the conventional monocentric model in which space is homogeneous by considering the impact of radial highways on which transport costs are lower. The result is a star-shaped city where each of the 'fingers' grows along a highway. Baum-Snow (2007a) was the first to provide convincing empirical evidence that highways caused suburbanization, as predicted by this model. Later studies have considered the impact of highways on the growth of metropolitan areas and employment (Duranton and Turner, 2012), as well as on increasing economic performance in peripheral areas (Banerjee et al., 2012; Chandra and Thompson, 2000; Michaels, 2008).

The analysis of the effects of highways on population size in peripheral areas has been studied using a different identification strategy, according to which the connection to a highway network of a peripheral town is regarded as an unintended consequence of its location between larger urban areas. Therefore, highway assignment was considered as random within the set of possible trajectories that connect the largest cities (Banerjee et al., 2012; Chandra and Thompson, 2000; Fajgelbaum and Redding, 2014; Michaels, 2008).

We follow the literature and assume a monocentric model setting of a closed city, in which homogeneous workers commute to the city center where employment is concentrated.[2] Building height and residential location are chosen by workers under utility maximization conditions. Rents and population density decrease with distance from the city center. We distinguish between three area types: (i) The central city, closest to the CBD, (ii) the suburb, between the central city and the edge of the city, and (iii) the area outside the city, which is called the periphery, in which population density is exogenously given.

Improvements in the highway network reduce commuting costs per unit distance. Workers then choose a residential location further away from the city center, which reduces population density within the central city (Baum-Snow, 2007b), and increases population in areas which are connected to the highway

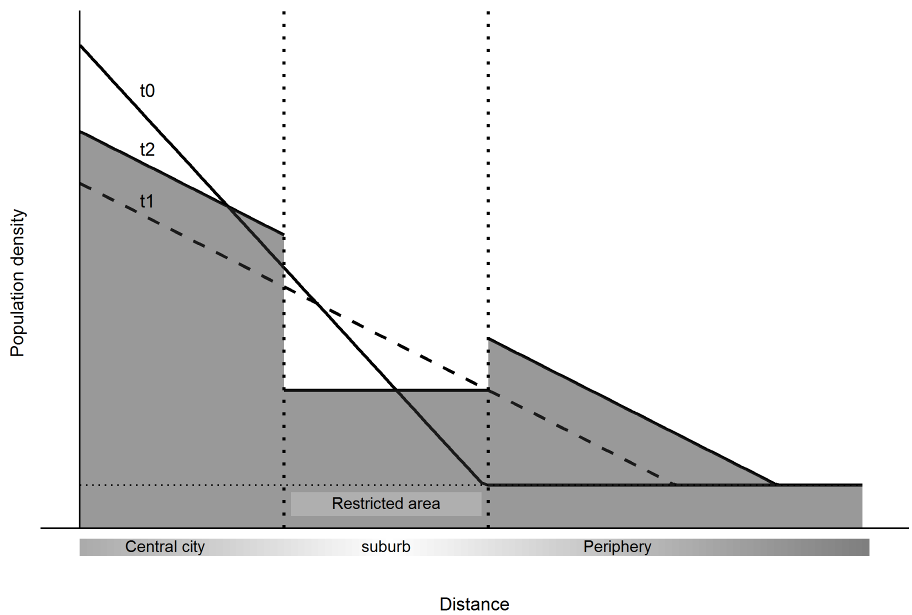
[2] The main consequence is that we assume that total population is not affected by land policies.

network.[3]

These effects may change when land policies are introduced. Figure 1 illustrates the effects of highways given land development restrictions in designated areas in the outskirts of the city. When no highways or land development restrictions are present, population density decreases in a relatively quick rate with the distance from the city center, until the city edge (t_0). The construction of a new highway will reduce commuting costs and result in outward expansion of the city edge (see t_1), but development restrictions would result in a decline in population density in restricted areas, and population redistribution to adjacent areas (see t_2).

The model predicts that the combined effect on population growth in the central city and suburb areas may be either positive or negative. This depends on the magniude of the positive effect following an increase in highway extent and decrease in commuting costs, as well as on the strictness of the development and its subsequent negative effect on growth. However, for population in peripheral areas the effects of highways and planning restrictions are both positive, as population density is expected to increase due to the reduction in commuting costs and the redistribution of population from the restricted area. The combined outcome reflects a leapfrog pattern of low population growth rates in the suburbs, but high growth rates in peripheral areas.

Figure 1: Illustration of the theoretical framework



Notes: t_0: No highways, no development restrictions. t_1: Highways, no development restrictions. t_2: Highways and development restrictions.

[3]Highway improvements may also increase the attractiveness of an urban area, and its ability to attract new workers (Duranton and Turner, 2012).

3. Spatial planning and transportation development in the Netherlands

3.1. Land development restrictions

The origins of Dutch land use planning date back to the early 20th century when the Housing Law (1901) obliged cities to make a plan for large scale extensions of their residential areas. Areas without such plan could not be developed, at least not on a large scale. However, the planning system was universally implemented after World War II. During the 1950's it became clear to policy makers that the population should be expected to continue to grow in the next decades and that this would have substantial consequences for land use. In line with prevalent ideas about government intervention in the economic system, it was thought that land use planning could contribute to an orderly development of national land use that would increase social welfare.[4]

The 1958 document on 'The Development of the Western Part of the Country' (*Rijksdienst voor het Nationale Plan*, 1958) presents a vision on spatial planning of the Randstad that would remain dominant until the 1980s. "If one allows the current development to proceed, one of the main advantages of the Dutch Randstad in comparison to foreign conurbations will be forfeited: the spatially separately located cities of transparent size". This vision was translated into several main policy measures. First, the center of the urbanized Randstad, the 'Green Heart', was preserved for agricultural use (Koomen et al., 2008; Koomen and Dekkers, 2013). Second, concerns that expansion of large urban areas would result in a formation of one large urban agglomeration were addressed by assignment of 'buffer zones', areas surrounding large cities. Spatial delineation of the zones appeared later in 1966 (Dieleman et al., 1999; Koomen et al., 2008). Third, the central government was also involved in directing urban growth. Areas outside of main cities were defined as growth cities (*'Groeikernen'*), which were destined to absorb suburbanization. The definition of the growth cities and the implementation of the new policy was not fully realized before the late 1970s and early 1980s (following the Third Report on Physical planning, 1974–1977),[5] and was eventually discontinued by the early 1990's when new national policy redirected urban development back to areas in vicinity to traditional city centers (Geurs and Van Wee, 2006; Jobse et al., 1991; Ostendorf and Musterd, 1996).[6]

3.2. Urban and road network development

Intercity passenger and freight transportation was historically based on railway and inland waterways. The first railway line was opened in 1839, and the railway network reached its peak length in 1930 (Koopmans et al., 2012).

[4]Spatial planning in the Netherlands is guided by a series of white papers, or Reports on Physical Planning (Nota Ruimtelijke Ordening, released in 1960, 1966, 1974-1977, 1988, and 2001).

[5]The designation of the growth cities (*'groeikernen'*) was soon followed by designation of four growth cities (*'groeisteden'*), existing cities located far from population concentration which were destined to absorb additional urban population growth. We do not make a distinction between groeikernen and groeisteden.

[6]This was directed by the Fourth Report on Physical Planning in the Netherlands (*Vierde Nota Ruimtelijke Ordening*, 1989), and particularly its annex report in 1991 (*Vierde Nota Ruimtelijke Ordening Extra*, abbreviated as 'VINEX').

Motorized transport was quickly adopted after the introduction of automobiles in 1896 (Smaal, 2012). By 1930 there were approximately 68,000 registered private auto-mobiles, 30,000 motorcycles and 44,000 busses and trucks (Veenendaal, 1996). The main roads were then built based on a plan which was first laid out in 1821 (see Appendix Appendix A). In 1927 the government laid out its first official road network plan (*'Rijkswegenplan'*).[7] The first highway (between The Hague and Utrecht) was opened in 1937. By 1960 the network reached a length of 351 km (Smaal, 2012). During the same period, between 1946 and 1960, car ownership grew from 5 cars per 1,000 people, to 45.8 cars per 1,000 people (Ligtermoet, 1990), still only a fraction of contemporary car ownership rates in the United States.[8]

Following growth in traffic, the government decided in 1961 to expand the network, and to construct additional 1,200 kilometers of highway before 1975. This goal was reached in 1972 (Ligtermoet, 1990). Car ownership rates also grew rapidly to 218 cars per 1,000 people, approximately half of US car ownership rates at the time. After the completion of the expansion plan, funds for road investments were exhausted. Around the same time, increasing attention to environmental impacts and congestion effects of car travel, and the breaking of the oil crisis in 1973 (after which driving on Sunday was prohibited for two months), led to a new government policy which aimed to decrease the dependency in commuting by car (Ligtermoet, 1990; Schwanen et al., 2004; Smaal, 2012). The expansion of the highway network during the 1960s is now regarded as an outdated policy. The development of the highway network continued at a slower after the 1970s,[9] and it was characterized by the debate between the demand for better roads and environmental preservation and limitation of energy usage.

4. Data

We estimate the effect of new highways on population growth distinguishing between (i) central cities (ii) suburban municipalities and (iii) peripheral municipalities. Growth of the highway network occurred primarily between 1961 and 1972. Hence, we will examine population growth just after this period, so between 1970 and 1980. The choice of the exact time period was made to maintain consistency with highway variables, which are only available for ten years intervals (1960, 1970 and 1980). As a sensitivity analysis we also analyze population growth between 1980 and 1990 to examine the long-term effects.

We make use of historical data on the extent of transportation networks and population, calculated for 811 municipalities for 1980 municipal boundaries (using 250 square meter cells). Table 1 provides a descriptive summary of the variables used in the analysis.

Our main dependent variable is population growth between 1970–1980 per municipality, which was obtained from Statistics Netherlands. On average, population grew by 22 percent between 1970 and 1980. Data on the 1960, 1970

[7] This plan was revised several additional times in the following decades, to consider updated projections of traffic and to specify different road capacity types.

[8] By comparison, in 1960 there were 411 cars per 1,000 people in the United States.

[9] See highway network expansion maps in Appendix Appendix A.

Table 1: Summary of main variables

Statistic	Year	Mean	St. Dev.	Min	Max
Pop. growth (1970-1980)	1970-1980	0.199	0.254	−0.641	2.927
Pop. growth (1980-1990)	1980-1990	0.068	0.151	−0.5606	1.94
Highway rays	1970	1.321	1.367	0	8
Highway density	1970	68.7	113.5	0	941.7
Rail stations	1930	1.02	1.748	0	18
Population (1930)	1930	9,718.10	38,786.10	0	743,900.7
Population (1960)	1960	14,005.7	48,420.5	34.797	855,539.7
Buffer zone share	1966	0.046	0.182	0	1
Nature coverage share	1960	0.098	0.139	0	0.921
Green heart share	1958	0.148	0.353	0	1
Distance from central city	1980	22.691	15.362	0	95.433
Reclaimed land share	1980	0.014	0.098	0	1
Growth city	1977	0.022	0.147	0	1
Road density (instrument)	1821	54.3	101.8	0	727.4
Artificial buffer zone share (instrument)	1980	0.212	0.409	0	1

Notes: (i) The number of observations is 811. (ii) Highway and road density are defined as length in meters per square km. (iii) Municipalities with zero population in 1930 are in areas which were reclaimed from the sea. (iv) We report population growth (1970–1980) excluding two municipalities, which were built on land reclaimed from the sea. (v) We report population growth (1980–1990) excluding one municipality (Almere), which was built on land reclaimed from the sea.

and 1980 highway network was obtained through the Historisch NWB (*Nationaal Wegenbestand*).[10] Information regarding the main roads in 1821 was available through the ministry of infrastructure and environment (see appendix Appendix A). For 1970, road data was used to create two variables that measure highway extent: highway density and rays. The average distance to highway access point was also used as an alternative measure of highway extent. The results of this measure produce similar coefficients, but are less trustworthy, as shown by a low first-stage Kleibergen-Paap F-test score which implies a weak instrument. Highway access points are frequent in the Netherlands, and access to highways is almost continuous. In 1970, the highway network had approximately 340 access points, which corresponds with an access point for every 2.8 highway kilometers on average.

Highway density is calculated as the ratio of the meter length of highways in a municipality and the municipal area (in square kilometer). We calculate highway rays from each municipal area center following Baum-Snow (2007a) and Baum-Snow et al. (2017). We define a 5 kilometers radius around each municipal centroids (defined by neighborhood population weights), and count the number of times highways cross this radius.[11] Table 2 provides a descriptive summary of both highway extent variables. Our measures of highway extent complement each other. The use of highway rays to study the effect of highways on suburbanization is common in literature. However, highway density may better reflect highway accessibility in rural areas (particularly if access points are

[10] Highways were identified based on road type indication of dual motorway or highway, or whether the road is maintained by the central government. Both result in similar figures of highway network length, as reported in Ligtermoet (1990) and Statistics Netherlands (2015).

[11] Radius of 5 kilometers was determined based on common municipality areas. A sensitivity analysis included using rays based on 3 kilometers radius, which have shown little differences in coefficient values and statistical significance of the estimators.

frequently present), or in municipalities with large areas or irregular boundary shapes, in which highways cross the municipal area but do not directly reach population centers.[12] We use both highway extent measures in order to reconcile such possible differences in measurements. We define central cities as cities with a population exceeding 50,000 people in 1930.[13] Since municipalities vary in area, we also restricted the definition of central city to municipalities which had population density level of at least 500 inhabitants per square kilometers in 1930.[14] Suburban municipalities are defined as municipalities directly adjacent to central cities, or located within five kilometers from the centroid of a central city.[15] This definition is in line with the idea that these municipalities are most likely to experience population expansion following improvement in highways. All other municipalities are defined as peripheral municipalities. In total, we define 20 central cities, 133 suburban municipalities and 658 peripheral municipalities (Figure 2). The use of historical population levels to define municipality types is also used in order to relief suspicion of endogeneity in the assignment of central cities, suburbs and peripheral municipalities.

Suburban municipalities face development restrictions when they are located within the Green Heart or within buffer zones. Buffer zones cover 12.9 percent of suburban municipalities' area, compared to 7.9 and 2.6 percent of central cities and peripheral municipality's area. The green heart covers 17.4 percent of suburban municipalities' area, compared to 4.6 and 13.8 percent of central cities and peripheral municipalities' area.[16]

We also use information on rail stations as a control variable. Because railway length reached its maximum in 1930 it is possible that the presence of stations in 1930 interfered with the effects of new highways. Data on historical railway stations was obtained from Koopmans et al. (2012). Data on nature coverage in 1960 was available from Alterra (Kramer, 2005). Spatial planning and land development restriction data, including the boundary of the green heart and the buffer zones, was obtained from Koomen et al. (2008).

The Netherlands is characterized by abundance of water, which often obstructed land development. During the 20th century large land reclamation projects were carried out in the Netherlands, and such lands were often designated for development purposes. To control for population growth in land reclaimed from water we include the share of municipal area which was reclaimed between 1930-1980, which was calculated by comparing water line boundaries

[12] Peripheral municipalities often include several small villages, and therefore can have multiple town centers.

[13] Appendix Appendix B.1 includes a sensitivity analysis with population thresholds of 35,000, and 65,000 inhabitants.

[14] Despite having a population of approximately 60,000 inhabitants in 1930, the municipality of Apeldoorn was not included as a central city. Apeldoorn has the largest municipal territory in the Netherlands with 339 square kilometers of municipal area, of which 81% is open space. With a population density of 176 people per square km in 1930, it is rural in character.

[15] We also tested suburb definitions of municipalities located 10, 15 and 20 kilometers from the centroid of central cities. Highway extent coefficients maintain relatively similar values and their statistical significance level, with the exception of rays under the 20 kilometer radius suburbs scenario, where the effect becomes statistically insignificant.

[16] The correlation between the share of municipal area included within the buffer zones and the suburb municipalities dummy is 0.209. The correlation between the share of area included within the Green heart and suburb municipalities dummy is 0.355.

between 1930 and 1980.[17]

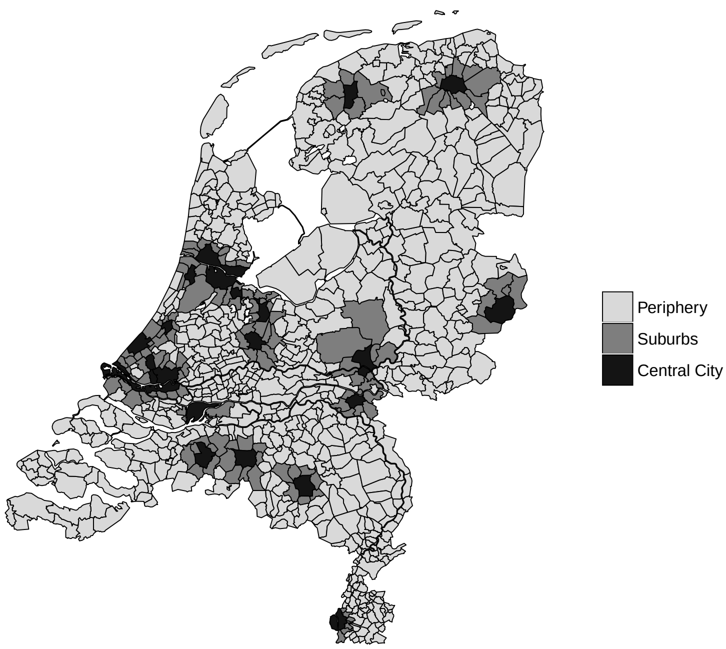
To guide urban sprawl away from areas designated to remain in natural or agricultural use, the Dutch government defined "growth cities" in which urban development was concentrated. The growth cities lie within eighteen 1980 municipalities, of which two are central cities (Breda, Groningen), 6 are suburban municipalities and 10 are peripheral municipalities. The delineation of the growth cities was implemented in the late 1970's.

Table 2: Summary of highway variables

	Municipality group	N	Mean	St. Dev.	Min	Max
Highway rays	Central city	20	2.65	1.182	1	5
	Suburb	133	1.789	1.446	0	7
	Periphery	658	1.185	1.315	0	8
Highway density	Central city	20	174.1	98	0	423.1
	Suburb	133	75.6	111.8	0	584.6
	Periphery	658	64.2	112.8	0	941.7

Note: Highway density is defined in meters per square km.

Figure 2: Municipalities (1980)



[17]The 1980-1990 analysis includes the same variable, adjusted to 1990 levels.

5. Methodology

5.1. Main specification: Population growth (1970–1980)

We investigate population growth in urban areas by specifying the following specification:

$$\begin{aligned}\Delta \ln(Pop_i) =& \beta_0 + \beta_S S_i + \beta_C C_i + \beta_{HP} H_i P_i + \beta_{HS} H_i S_i + \beta_{HC} H_i C_i \\ &+ \sum_k \alpha_k X_{i,k} + \epsilon_i,\end{aligned} \tag{1}$$

where: $\Delta \ln(Pop_i)$ is the change in log population in municipality i. P_i indicates a peripheral municipality, S_i indicates a suburban municipality, and C_i indicates a central city. H_i defines highway extent, highway rays or highway density (in log). H_i is interacted with P_i, S_i and C_i to allow distinct effects of highways by type of municipality. $X_{i,k}$ refers to control variables, including the number of rail stations and population level, both in 1930, nature coverage in 1960, the municipality share within a buffer zone, and the municipality share within the green heart. We will also estimate another specification in which we restrict $\beta_C = \beta_S$ and $\beta_{HC} = \beta_{HS}$. This restricted specification allows us to study the effects of highways in urban agglomerations as a whole, consisting of central cities and suburbs.

We address the concern that the development of the highway network is most likely endogenous, as its construction is likely influenced by travel demand. We use an instrumental approach using the 1821 road network as an instrument (see also Baum-Snow (2007a,b); Baum-Snow et al. (2017); Duranton and Turner (2012); Garcia-López (2012); Garcia-López et al. (2015); Pasidis et al. (2015).[18]

Many spatial development restrictions and land use policies in the 1960s and 1970s are also suspected to be endogenous. This is unlikely the case for buffer zones, because their definition relies on their exogenous location in immediate adjacency to traditional urban centers. Hence, in our main analysis we consider buffer zone share as exogenous. Nevertheless, in a sensitivity analysis we consider them as endogenously determined, employing an (artificial) buffer area of fifteen kilometers around central cities as an instrument. We find the results unchanged for the impact of highways. We also take into account that the effect of highways may differ for areas within and outside buffer zones, and use additional interaction terms between highways, buffer zones and municipality types.

Our sample includes eighteen growth cities that were designated late during our study period (1977), and are likely to be dependent on the development of highway accessibility at the time. We therefore do not control for the presence of growth cities. Our results are also robust when these eighteen municipalities are excluded from the analysis.

[18] The road network system of 1821 was created before the industrial revolution and a century of rail-transport dependency, and thus unlikely to have been affected by the changes in spatial structure that resulted from improvements in transportation technologies in the following century. This argument suggests that rail infrastructure may also be used as a valid instrument for the highway network, as planned highway trajectories often followed existing historic railway lines, particularly where bridge crossing were already present. However, due to its close association with later population growth, it can be argued that rail accessibility does not satisfy the exclusion restriction.

5.2. Endogeneity issues of the interaction terms

The above specification includes interactions between highways extent variables and three municipality types: central cities, suburbs and peripheral municipalities. It follows that the three interaction terms H_iP_i, H_iS_i, H_iC_i may be considered as endogenous. In the context of an endogenous variable with interactions, several estimation procedures are then possible.

The standard approach, which we will refer to as the multiple first-stage approach, is to estimate a separate first-stage regression for each of the endogenous interaction variables.[19] A second approach, which we will show to be more efficient, is to estimate a single first-stage to predict $\hat{H}_i$ using an instrument, denoted by Z_i, and to use this predicted variable interacted with the dummies for central cities, suburbs and peripheral municipalities in the second stage (as suggested by Balli and Sorenson, 2013). Hence, one estimates the following first-stage equation:

$$H_i = \delta_0 + \delta_S S_i + \delta_C C_i + \delta_Z Z_i + \sum_k \gamma_k X_{i,k} + u_i, \tag{2}$$

where δ and γ_k are coefficients to be estimated. Given the first-stage estimates of the coefficients in (2), $\hat{H}_i$ is predicted, and is then interacted with P_i, S_i and C_i respectively. The resulting interaction terms $(\widehat{H_iP_i}, \widehat{H_iS_i}, \widehat{H_iC_i})$ are then used in a second stage. Robust standard errors in the second stage can be calculated following Angrist and Pischke (2008).

Because P_i, S_i, and C_i are mutually exclusive categorical variables (i.e, dummy variables that sum to one), the approach generates consistent and efficient estimators. To demonstrate this, first consider a multiple first-stage approach in which all coefficients are allowed to vary for periphery, suburbs or central cities:

$$P_iH_i = \delta_P P_i + \delta_{P,Z} P_i Z_i + \textstyle\sum_k \gamma_{P,k} P_i X_{i,k} + u_{p,i} \tag{3a}$$
$$S_iH_i = \delta_S S_i + \delta_{S,Z} S_i Z_i + \textstyle\sum_k \gamma_{S,k} S_i X_{i,k} + u_{s,i} \tag{3b}$$
$$C_iH_i = \delta_C C_i + \delta_{C,Z} C_i Z_i + \textstyle\sum_k \gamma_{C,k} C_i X_{i,k} + u_{c,i} \tag{3c}$$

where $P_i + S_i + C_i = 1$. It is well known that using (1) and (3) together generates consistent estimators (Wooldridge, 2002, p. 122). These first stages can equivalently be estimated jointly by summing equations (3a), (3b) and (3c):

$$\begin{aligned} H_i = {} & \delta_P P_i + \delta_S S_i + \delta_C C_i + (\delta_{P,Z} P_i + \delta_{S,Z} S_i + \delta_{C,Z} C_i) Z_i \\ & + \sum_k (\gamma_{P,k} P_i + \gamma_{S,k} S_i + \gamma_{C,k} C_i) X_{i,k} + u_i, \end{aligned} \tag{4}$$

where $u_i = u_{p,i} + u_{s,i} + u_{c,i}$. Hence, the single first-stage approach based on (2) is a special case of the multiple-first stage approach, given the restriction $\delta_Z = \delta_{P,Z}P_i + \delta_{S,Z}S_i + \delta_{C,Z}C_i$ and $\gamma_k = \gamma_{P,k}P_i + \gamma_{S,k}S_i + \gamma_{C,k}C_i$ for all $k = 1 \dots K$. This restriction can be tested using a standard F-test. If it holds, the single first-stage approach is more efficient than the multiple first-stage approach, as follows from a general econometric argument that imposing valid restrictions

[19] See Wooldridge (2002), page 122.

improves the efficiency of estimators. Furthermore, note that the single first-stage approach is less likely to suffer from weak instruments (as it avoids using additional instruments for each endogenous interaction term). Therefore there may be cases where the approach is not only more efficient, but it is also more likely to produce less biased results.

While both the multiple and single-first stage approaches assume that endogeneity in H_i results in endogeneity of the interaction terms, we also examine a third approach which treats the interaction terms as exogenous, under certain conditions. Bun and Harrison (2014) show that an interaction term between endogenous and exogenous variables may be treated as exogenous if the conditional expectation of the endogenous variable and the error term ($H_i u_i$) does not depend on the exogenous regressor with which the endogenous variable is interacted (P_i, S_i, C_i). The validity of this assumption can be tested using a standard Hausman test (Hausman, 1978), or an extended version which also tests the presence of weak instruments (Bun and Harrison, 2014; Hahn et al., 2011). This test compares the results of the model where the interaction term is treated as endogenous with the model in which it is treated as exogenous.[20]

We will apply all three approaches. It appears that these approaches result in similar coefficient values. The single first-stage approach produces the smallest standard errors and the highest first stage Kleibergen-Paap F-test value, and is therefore preferred.

5.3. Peripheral municipalities subsample

The literature on the effects of highways in peripheral regions commonly treats the assignment of highways as exogenous, as it is argued that highway assignment depends on the exogenous location between two larger population centers that are connected with a highway.[21] Plausibly, this exogeneity assumption does not hold in our analysis. Due to the relatively small spatial scale, it is likely that highway assignment in peripheral areas is endogenous. Trajectories of planned highways may have been directed to pass through faster growing peripheral towns, where the spatial scale is sufficiently small, such network planning decisions can be done without imposing costly road bypasses. To study the effects of highways in peripheral municipalities we estimate (1) on a restricted subsample of the peripheral municipalities. We will instrument highway extent as in (4).

6. Estimation results

6.1. Main results: Population growth (1970–1980)

To apply the single first-stage approach, we first test whether the restrictions imposed are valid using a standard F-test. The results of the hypothesis test

[20] Following Bun and Harrison (2014), we also use Wald test to test the hypothesis of consistency of OLS estimators, in which all variables are considered exogenous. This hypothesis is rejected.

[21] See, for example, Chandra and Thompson (2000); Fajgelbaum and Redding (2014); Michaels (2008).

of these restrictions show that for both highway measures we cannot reject the null hypothesis, which indicates that the single first stage approach is valid.[22]

The estimation results of model (1) using the single first-stage approach, are presented in Table 3.[23] The results in columns 1 and 2 show that one highway ray increases municipal population growth by 15.9%. One percentage increase in highway density is expected to increase population growth by 0.058%. This corresponds with an increase of approximately 10% given one standard deviation increase in highway density. The effect of highway rays and density on population growth in central cities and suburbs is small and statistically insignificant. The effects of buffer zone are negative and significant. An increase in 1% in share of municipal area defined as a buffer zone is expected to drop population growth by approximately 12.7-14.9%. This confirms the findings of Geurs and Van Wee (2006); Koomen et al. (2008); Koomen and Dekkers (2013) that the Dutch policy of open space preservation was effective in preventing urban sprawl in designated areas. A positive coefficient for suburbs and central cities (column 2) indicates that urban municipalities experienced stronger population growth compared with peripheral areas.

The results of the restricted specification (columns 3 and 4), in which we assume an identical effect of highways for agglomerated urban areas (central cities and suburbs), also show positive effects of highways on peripheral population growth. The effect of increase in highway rays on population growth is estimated at 17.2%. The effect of one percentage increase in highway density is essentially also the same, at 0.064%. We find a significant effect of highway density on population in urban agglomerations (central cities and suburbs, combined), but this effect disappears when highway rays are examined.

The result that highways have hardly any effect on central city population, but a strong effect on the peripheral municipalities, largely confirms the findings of Baum-Snow (2007a,b) and related papers. However, the absence of a clear effect on population growth in suburban municipalities is not in line with previous findings. A possible explanation is the small spatial scale in the Netherlands. It is possible that municipalities adjacent to central cities do not experience a significant reduction in commuting costs following the construction of a new highway, and commuters from these municipalities still prefer local urban roads, or public urban transportation systems. Moreover, it may be that commuting costs remain relatively low in municipalities located further away from central cities, in that sense, the strong positive effect of highways found in peripheral municipalities could be interpreted as reflecting this suburbanization effect.[24]

We estimate an additional specification of the model in which we consider buffer zone share to be endogenous. Here we instrument buffer zones using an 'artificial buffer' variable, defined by a dummy which indicates municipalities

[22] Both F-tests show values lower than 0.8, so with a corresponding p-value exceeding 0.7. The F-test has 17 degrees of freedom. We restrict the nine coefficients of highway extent (δ_Z) and eight control variables (γ_k) to be equal between municipality types. Because distance to central cities has no variation within central cities, we have 17 restriction.

[23] First stage estimation results show a positive and statistically significant effect of the instrument on highway extent in the 1970s, see Table C.5.

[24] Twelve suburban municipalities are adjacent to two central cities. We have also considered a specification in which these suburban municipalities experience a double effect. The results maintain very similar values, and as expected, the estimated effect of $H_i S_i$ becomes slightly smaller.

Table 3: Population growth and highways – Main results

Dependent variable: Log. pop. growth (1970–1980)						
	Main model		Restricted specification		Main model	
	(1)	(2)	(3)	(4)	(5)	(6)
Highway rays*periphery	0.15982***		0.17245***		0.15262***	
	(0.06184)		(0.05843)		(0.05596)	
Highway rays*suburb	−0.00456		−0.03959		−0.00780	
	(0.07982)		(0.06623)		(0.06701)	
Highway rays*central city	−0.01132		−0.03959		−0.08769	
	(0.10705)		(0.06623)		(0.11379)	
Highway density*periphery		0.05877***		0.06461***		0.05638**
		(0.02155)		(0.02050)		(0.02393)
Highway density*suburb		−0.02957		−0.04516*		−0.01423
		(0.02980)		(0.02396)		(0.02266)
Highway density*central city		−0.03991		−0.04516*		−0.04079
		(0.02919)		(0.02396)		(0.03322)
Suburb	0.28349	0.21933*	0.37373***	0.29253***	0.29345**	0.17887**
	(0.18291)	(0.11396)	(0.13702)	(0.09324)	(0.13192)	(0.07812)
Central city	0.16184	0.11401	0.37373***	0.29253***	0.34126	0.11201
	(0.26903)	(0.14813)	(0.13702)	(0.09324)	(0.29880)	(0.16357)
Buffer zone share	−0.14900***	−0.12744***	−0.13571**	−0.12089***	−0.38633**	−0.18645
	(0.05225)	(0.03978)	(0.05264)	(0.04071)	(0.17001)	(0.13901)
Rail stations	0.00238	0.00152	0.00110	0.00006	0.00119	0.00095
	(0.00345)	(0.00354)	(0.00357)	(0.00369)	(0.00338)	(0.00362)
Log Pop. (1930)	0.00519	−0.00658	0.00374	−0.00938	−0.00123	−0.00790
	(0.03940)	(0.03624)	(0.03965)	(0.03604)	(0.03972)	(0.03855)
Log Pop. (1960)	−0.05521	−0.04869	−0.05821	−0.05421	−0.05175	−0.04844
	(0.04862)	(0.04638)	(0.04798)	(0.04599)	(0.04909)	(0.05004)
Nature coverage share	−0.06932	−0.04736	−0.06344	−0.03564	−0.09461**	−0.05597
	(0.04530)	(0.05008)	(0.04465)	(0.04996)	(0.04719)	(0.05173)
Green heart share	−0.03986	0.02458	−0.03631	0.03189	−0.04069	0.02200
	(0.03806)	(0.02279)	(0.03880)	(0.02351)	(0.03893)	(0.02386)
Distance from central city	−0.00125	−0.05355***	0.01677	−0.03071**	−0.01464	−0.05682***
	(0.01876)	(0.01374)	(0.01560)	(0.01258)	(0.01781)	(0.01465)
Reclaimed land share	0.60556*	0.52545	0.58282*	0.49228	0.57074	0.52261
	(0.34107)	(0.32010)	(0.34253)	(0.31686)	(0.34757)	(0.33426)
Constant	0.41606***	0.68340***	0.38200***	0.66955***	0.49993***	0.71033***
	(0.09289)	(0.13068)	(0.09197)	(0.13131)	(0.10734)	(0.13993)
Highways instrumented	Yes	Yes	Yes	Yes	Yes	Yes
Buffer zone instrumented	No	No	No	No	Yes	Yes
Kleibergen-Paap F-statistic (Highways)	10.96	33.62	10.97	32.50	7.994	13.75
Kleibergen-Paap F-statistic (Buffer zone)					16.15	16.15
Observations	811	811	811	811	811	811

Notes: (i) Highway extent variables are in 1970 levels. (ii) Highway density is expressed in logarithm, (iii) Possible endogeneity in the interaction terms is addressed following the single first-stage approach (Balli and Sorensen, 2013). (iv) Robust standard errors in parentheses. $p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$

that are completely contained within a radius of 15km from the cities around which buffer zones were present (Amsterdam, Rotterdam, The Hague, Utrecht and Maastricht).[25] The estimation results (Table 3, columns 5-6) show that the effect of highways in peripheral municipalities remains robust when we consider buffer zones to be endogenous.[26]

6.2. Endogeneity in the highways interaction terms

In Table 4 we present the results of other approaches to address endogeneity in the interaction terms. We compare the results of the main model as presented in Table 3 using (i) OLS, (ii) assuming that the interaction terms are exogenous (following the approach of Bun and Harrison, 2014), and (iii) assuming endogenous interaction terms, and estimating separate first steps for each term as commonly applied (e.g, Wooldridge, 2002). Note that the Kleibergen-Paap F-test of the single first-stage approach (presented in Table 3) is higher than the values of the test for exogenous interaction terms and multiple first-stage approaches.

The results show that for all instrumented approaches, and for both highway measures, the estimated interaction coefficient with peripheral municipalities obtains similar values. Note that the multiple first-stage approach for highway density (columns 3 and 6) are invalid, as demonstrated by a low F-test value, implying that the results based on this approach are biased.

Notably, the effects of highways are found to be substantially lower in value when estimated in OLS (columns 1 and 4) compared with the instrumented estimations. Since reversed causality is controlled, and would be expected to result in overestimation rather than underestimation of the effects, the lower coefficient value is possibly related to omitted variable bias. For instance, it may be that highway planners initially decided that highway trajectory would reach 'growing' population centers, but without entering them, in order to allow highway access while avoiding expected negative highway noise externalities.

6.3. Peripheral municipalities subsample

The results using a subsample of peripheral municipalities are presented in the Appendix B.2, Table B.3. They show similar effects as in Table 3. Since the previous literature generally considers highways in peripheral areas as exogenous, we also estimate the restricted specification using OLS. The results show that the effects of highways become substantially smaller. The effect of an increase in one highway ray is about 1.38%, and the effect of 1% increase in highway density is about 0.0105%, indicating a negative bias. This implies that in contrast with previous assumptions in other studies, highway extent in peripheral areas in the Netherlands cannot be regarded as exogenous.[27]

[25] We also experimented with artificial buffers in ranges between 3-20 kilometers radius from central cities. A buffer of fifteen kilometers was chosen as it is in line with actual buffer zones ranges.

[26] A standard Hausman test for the endogeneity of buffer zone share provides a value of 1.91, corresponding with a p-value of 0.166 which suggests that the exogeneity assumption is not rejected.

[27] Excluding eighteen observations that refer to growth cities generates almost identical results.

Table 4: Analysis of endogeneity of interaction terms

Dependent variable: Log. pop. growth (1970-1980)						
	OLS	Exogenous interaction terms	Multiple first-stage	OLS	Exogenous interaction terms	Multiple first-stage
	(1)	(2)	(3)	(4)	(5)	(6)
Highway rays*periphery	0.01593***	0.12191**	0.12674**			
	(0.00435)	(0.05429)	(0.05952)			
Highway rays*suburb	0.00384	0.00868	0.12491			
	(0.01678)	(0.01738)	(0.14907)			
Highway rays*central city	−0.01301	−0.01122	−0.31745			
	(0.01383)	(0.01430)	(1.07391)			
Highway density*periphery				0.01228***	0.04115***	0.04565
				(0.00288)	(0.01535)	(0.11888)
Highway density*suburb				−0.00163	0.00012	0.04594
				(0.00897)	(0.00878)	(0.45221)
Highway density*central city				−0.01271*	−0.01296**	−5.49310
				(0.00676)	(0.00656)	(275.74290)
Suburbs	0.03703	0.18742**	0.00169	0.04907	0.10486*	−0.00718
	(0.02668)	(0.08289)	(0.21536)	(0.04139)	(0.05693)	(1.39078)
Central city	−0.18358***	0.02790	0.87568	−0.15349**	−0.06437	26.40456
	(0.06358)	(0.14063)	(2.71406)	(0.06341)	(0.08639)	(1,330.757)
Buffer zone share	−0.12058***	−0.13671***	−0.17805*	−0.11564***	−0.11559***	−0.00568
	(0.04012)	(0.04573)	(0.09083)	(0.03738)	(0.03908)	(6.46219)
Rail stations (1930)	0.00354	0.00382	0.00824	0.00305	0.00197	0.04551
	(0.00335)	(0.00461)	(0.01401)	(0.00338)	(0.00353)	(2.25659)
Log Pop. (1930)	−0.02194	−0.00029	0.00682	−0.02187	−0.01333	0.00246
	(0.03384)	(0.04037)	(0.04174)	(0.03286)	(0.03404)	(0.66244)
Log Pop. (1960)	−0.01436	−0.04393	−0.05765	−0.01758	−0.03712	−0.09695
	(0.03890)	(0.04755)	(0.05293)	(0.03719)	(0.04145)	(2.62362)
Nature coverage share	−0.06825	−0.05248	−0.07327	−0.06229	−0.04613	−0.14641
	(0.04453)	(0.05356)	(0.06851)	(0.04521)	(0.04967)	(4.26074)
Green heart	0.02011	−0.03407	−0.03946	0.02910	0.02933	0.01873
	(0.02296)	(0.04338)	(0.05314)	(0.02228)	(0.02308)	(0.03996)
Distance from central city	−0.04276***	−0.02120	−0.01385	−0.04780***	−0.05165***	−0.05590
	(0.01294)	(0.01983)	(0.02673)	(0.01319)	(0.01465)	(0.19903)
Reclaimed land share	0.48798	0.60396*	0.64925*	0.47617	0.49514	0.60541
	(0.31882)	(0.34187)	(0.34611)	(0.31352)	(0.30973)	(4.57343)
Constant	0.58241***	0.46841***	0.49801***	0.61804***	0.66816***	1.02210
	(0.11289)	(0.11834)	(0.13249)	(0.10879)	(0.12072)	(15.54208)
Instrumented	No	Yes	Yes	No	Yes	Yes
Kleibergen-Paap F-statistic		3.618	0.037		9.005	0.0001
Observations	811	811	811	811	811	811
R^2	0.1691			0.1755		

Notes: (i) Instrumented variables in columns 2 and 6: highways, interacted in the second stage with regional dummies. Instrumented variables in columns 3 and 7: highway*periphery. Instrumented variables in columns 4 and 8: highway*periphery, highway*suburbs, highway*central city. (ii) Highway extent variables are in 1970 levels. (iii) Highway density is in logarithm. (iv) Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01

7. Sensitivity analysis and long-term effects

7.1. Interactions with buffer zones

Our main model specification tests the effects of highways for several municipality types, controlling for presence of buffer zones. It may also be that the effects of highways vary with the presence of buffer zones. To test this we extend our model by including an additional set of interaction variables between highway extent and buffer zones.[28] First, we assume that buffer zone share is exogenous. We later test our results under an endogenous buffer zone share assumption. As before, our instrumental variable strategy follows the single first-stage approach to compute each of the three instruments (for the three endogenous variables - highway extent, buffer zone share and highways∗buffer zone share). The instruments are first estimated using first-stage regressions, and then interacted in the second-stage with the three municipality type dummies.

The results in Table 5 (columns 1 and 2) show positive, but weakly significant, effects of highways in buffer zones within peripheral and suburban municipalities. Highway density has a positive effect of 0.15 within suburban buffer zones, significant at the 10% level. The effect of highways in buffer zones within central cities is positive (and statistically significant). Note that this effect is estimated based on only 11 observations with non-zero buffer zone share, so is not reliable. Notably, the effect of highways in peripheral areas remains robust and maintains similar values (and significance levels) when highway and buffer zones interactions are included.

When the buffer zone share is considered endogenous (Table 5, columns 3 and 4), the effects of highway density in peripheral buffer zone becomes negative and statistically significant, with a value of -0.20. This implies that due to the presence of buffer zones, the positive effect of highways in peripheral areas is nullified.[29] This means that highways reduced population growth in peripheral buffer zones, compared with peripheral municipalities outside buffer zones, with similar highway density. The effect of the interactions of highway rays with buffer zone share in all municipality types is statistically insignificant, suggesting that all the effects are captured by the main effects. Furthermore, the effect of highways in peripheral municipalities remains robust. The effect of highway density in suburban municipalities is negative at the 10% level, weakly suggesting that highways have contributed to slower population growth rates in suburban areas, consistent with a leap-frog pattern.

7.2. Effects on population growth (developments after 1980)

We have also estimated our model to examine the effects of highways on population growth one decade later, between 1980-1990 (Table B.4). It appears that there is no statistically significant effect of highway extent in the full sample,

[28] We focus here on buffer zone share as it is found to have a strong negative effect on population growth. The other type of regulation, the green heart, did not have any effect on population growth and is therefore a less interesting variable to explore.

[29] The combined effect of highways in peripheral buffer zones is calculated as $0.0705 - 0.2078 = -0.137$. However, a standard F-test shows that we cannot reject the null hypothesis that this combined effect is zero.

Table 5: Population growth and highways - Effect of highways in restricted areas

(Dependent variable: Log. pop. growth (1970–1980))				
	Main model with interaction highways*buffer zones			
	Endogenous highways		Endogenous highways and buffer zone share	
	(1)	(2)	(3)	(4)
Highway rays*periphery	0.16111**		0.14408***	
	(0.06555)		(0.04438)	
Highway rays*suburb	−0.03784		0.00883	
	(0.09862)		(0.07580)	
Highway rays*central city	−0.15251		0.01536	
	(0.11119)		(0.15035)	
Highway rays*periphery*buffer	0.09540		0.43895	
	(0.13122)		(0.84960)	
Highway rays*suburb*buffer	0.29416		−0.33480	
	(0.19135)		(0.48195)	
Highway rays*central city*buffer	1.44828**		0.66940	
	(0.64644)		(1.32445)	
Highway density*periphery		0.05595***		0.07059***
		(0.02128)		(0.01917)
Highway density*suburb		−0.04437		−0.10010*
		(0.03264)		(0.05670)
Highway density*central city		−0.09755***		−0.00037
		(0.02281)		(0.02954)
Highway density*periphery*buffer		0.06949		−0.20789**
		(0.06575)		(0.10364)
Highway density*suburb*buffer		0.15013*		0.46608
		(0.08111)		(0.29378)
Highway density*central city*buffer		0.61490***		0.37623
		(0.16426)		(0.99265)
Suburbs	0.34050	0.25702**	0.24772**	0.41212***
	(0.22361)	(0.12100)	(0.11825)	(0.14166)
Central city	0.52024*	0.36907***	−0.10619	−0.32836
	(0.28313)	(0.11664)	(0.40012)	(0.42583)
Periphery*buffer	−0.28910	−0.16499	−1.31957	0.34247
	(0.20624)	(0.10998)	(1.75975)	(0.27970)
Suburb*buffer	−0.80361*	−0.59905**	0.47317	−1.50225**
	(0.41060)	(0.24148)	(1.04017)	(0.63132)
Central city*buffer	−3.95784**	−2.97854***	0.24209	1.15492
	(1.81640)	(0.80984)	(2.44797)	(0.85749)
Constant	0.42278***	0.67607***	0.59248***	0.62789***
	(0.09445)	(0.13112)	(0.21033)	(0.14379)
Highways instrumented	Yes	Yes	Yes	Yes
Buffer zone instrumented	No	No	Yes	Yes
Highways * buffer zone instrumented	No	No	Yes	Yes
Kleibergen-Paap F-statistic (Highways)	11.05	33.86	7.994	13.75
Kleibergen-Paap F-statistic (buffer)			16.15	16.15
Kleibergen-Paap F-statistic (Highways*buffer)			11.06	7.423
Observations	811	811	811	811

Notes: (i) Included control variables are identical to the variables included in Tables 4 and B.3. (ii) Highway extent variables are in 1970 levels. (iii) Highway density is expressed in logarithm. (iv) Robust standard errors in parentheses (v) $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

whereas the effects become weakly significant when we use the subsample of peripheral municipalities.[30]

Possible reasons for the absence of evidence for strong highways effects one decade later is that much of the redistribution of population might have already taken place before 1980, during the years immediately following the great expansion of the highway network. This would suggest that highway effects estimated on the 1970–1980 data represent a long-run effect of approximately 10 to 20 years, after which the effect on population redistribution stabilizes and a new equilibrium is reached.[31]

Our interpretation is that Dutch planning policies aimed to prevent the formation of a large urban conurbation and to preserve agricultural activities and nature. While this policy was effective in achieving its original objectives, it had additional consequences when the highway network was extended. The restrictions that were imposed were compensated by strong population growth in peripheral areas, which resulted in increased commuting distances and time (Schwanen et al., 2001, 2004). Spatial policies in the following decades addressed this by enforcing stricter development restrictions regarding the preservation of open space, and attempting to direct urban population and employment growth back to existing urban areas.[32]

8. Conclusion

There is a large literature which shows that new highways depopulate city centers. We examine this in the context of the Netherlands where land development restrictions are common. Our analysis focuses on the expansion of the highway network in the 1960's, and its effects on population growth in central cities, suburbs and peripheral areas. We have addressed endogeneity issues by using 1821 road data as an instrument, and employed several innovative approaches to deal with endogeneity in the interaction of highway measures and different types of municipalities (central cities, suburbs and peripheral). In contrast to the literature, our findings for the Netherlands show no effect of highways for central cities and suburban areas. This finding is in line with the idea that strict planning policies and land development restrictions strongly interfere with the effects of highways. In line with the literature, our results show that the rapid development of the Dutch highway network had a substantial effect on changes in the population distribution, and that highways accelerated population growth in peripheral areas by about 10-15 percent. Hence, our results imply that when land development is restricted in the surroundings of cities, new highways divert population growth to locations further away from central cities. The development of the highway network results then in a large scale sprawl, which skips the suburbs and 'leapfrogged' to peripheral towns.

[30] The effects of buffer zones and nature coverage remain negative and significant, indicating that the presence of development restrictions continues to determine population growth in this period as well.

[31] An additional explanation is that the period 1980-1990 was characterized by national policies which aimed to reduce private car dependency and promote awareness of road externalities. This is also reflected in deceleration in car ownership compared with previous decades (see Figure A.3).

[32] See discussion in Dieleman et al. (1999); Geurs and Van Wee (2006).

Acknowledgements

We would like to thank Eric Koomen, Hans Koster and Miquel-Angel Garcia-Lopez for their useful comments and advices. Financial support from ERC Advanced Grant OPTION (#246969) is gratefully acknowledged. We would also like to thank Changjoo Kim and the audience of the paper presentation in the 63rd North American Regional Science Conference for a constructive discussion on the paper.

References

Agrawal, A., Galasso, A., and Oettl, A. (2017). Roads and Innovation. *Review of Economics and Statistics*, 99(3):417–434.

Alonso, W. (1964). *Location and Land Use. Toward a General Theory of Land Rent.* Harvard University Press, Cambridge, Massachusetts.

Anas, A. and Moses, L. N. (1979). Mode choice, transport structure and urban land use. *Journal of Urban Economics*, 246:228–246.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton university press.

Balli, H. O. and Sorensen, B. E. (2013). Interaction Effects in Econometrics. *Empirical Economics*, 45(1):583–603.

Banerjee, A., Duflo, E., and Qian, N. (2012). On the road: Access to transportation infrastructure and economic growth in China. *National Bureau of Economic Research*, (w17897).

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

Baum-Snow, N. (2007b). Suburbanization and transportation in the monocentric model. *Journal of Urban Economics*, 62(3):405–423.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., and Zhang, Q. (2017). Roads, Railroads, and Decentralization of Chinese Cities. *The Review of Economics and Statistics*, 99(3):435–448.

Bun, M. J. and Harrison, T. D. (2014). OLS and IV Estimation of Regression Models Including Endogenous Interaction Terms. *UvA-Econometrics Working Papers*, (2014/02).

Burchfield, M., Overman, H. G., Puga, D., and Turner, M. A. (2006). Causes of sprawl: a portrait from space. *The Quarterly Journal of Economics*, 121(2):587–633.

Chandra, A. and Thompson, E. (2000). Does Public Infrastructure Affect Economic Activity? *Regional Science and Urban Economics*, 30(4):457–490.

Dieleman, F. M., Dijst, M. J., and Spit, T. (1999). Planning the Compact City: The Randstad Holland Experience. *European Planning Studies*, 7(5):605–621.

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

Fajgelbaum, P. and Redding, S. J. (2014). External Integration, Structural Transformation and Economic Development: Evidence from Argentina 1870-1914.

Garcia-López, M.-À. (2012). Urban Spatial Structure, Suburbanization and Transportation in Barcelona. *Journal of Urban Economics*, 72(2-3):176–190.

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

Garcia-López, M.-À., Solé-Ollé, A., and Viladecans-Marsal, E. (2014). Do Land Use Policies Follow Road Construction. *CESifo working papers*, (4672).

Geurs, K. T. and Van Wee, B. (2006). Ex-post Evaluation of Thirty Years of Compact Urban Development in the Netherlands. *Urban Studies*, 43(1):139–160.

Glaeser, E. L., Gyourko, J., and Saks, R. (2006). Urban growth and housing supply. *Journal of Economic Geography*, 6(1):71–89.

Glaeser, E. L. and Kahn, M. E. (2004). Sprawl and Urban Growth. *Population English Edition*, 4(04):2481–2527.

Hahn, J., Ham, J. C., and Moon, H. R. (2011). The Hausman Test and Weak Instruments. *Journal of Econometrics*, 160(2):289–299.

Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, 46(6):1251.

Hsieh, C. T. and Moretti, E. (2017). Housing Constraints and Spatial Misallocation. *National Bureau of Economic Research*, (w21154).

Jobse, R., Kruythoff, H., and Musterd, S. (1991). Groeikernen in een veranderend ruimtelijk en demografisch krachtenveld. In *Geografisch Tijdschrift*, volume 25, pages 148–160.

Jun, M. J. and Hur, J. W. (2001). Commuting costs of "leap-frog" new town development in Seoul. *Cities*, 18(3):151–158.

Kok, N., Monkkonen, P., and Quigley, J. M. (2014). Land use regulations and the value of land and housing: An intra-metropolitan analysis. *Journal of Urban Economics*, 81:136–148.

Koomen, E. and Dekkers, J. (2013). The Impact of Land-Use Policy on Urban Fringe Dynamics: Dutch Evidence and Prospects. In Malkinson, D., Benenson, I., and Czamanski, D., editors, *Modeling of Land-Use and Ecological Dynamics.*, pages 9–35. Springer, Berlin Heidelberg.

Koomen, E., Dekkers, J., and Van Dijk, T. (2008). Open-space preservation in the Netherlands: Planning, practice and prospects. *Land Use Policy*, 25(3):361–377.

Koopmans, C., Rietveld, P., and Huijg, A. (2012). An Accessibility Approach to Railways and Municipal Population Growth, 1840–1930. *Journal of Transport Geography*, 25:98–104.

Kramer, H. (2005). Dataset Historisch Grondgebruik Nederland [HGN (1960)].

Ligtermoet, D. M. (1990). Beleid en planning in de wegenbouw : de relatie tussen beleidsvorming en planning in de geschiedenis van de aanleg en verbetering van rijkswegen. Technical report, Ministerie van Verkeer en Waterstaat.

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

Mills, E. S. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Area. *The American Economic Review*, 57:197–210.

Muth, R. F. (1969). *Cities and housing: the spatial pattern of urban residential land use*. University of Chicago Press.

Ostendorf, W. J. M. and Musterd, S. (1996). Groeikernen en compacte steden. Nieuwe grenzen voor de ruimtelijke ordening. *Planologisch Nieuws*, 16(2):91–101.

Pasidis, I., Viladecans-Marsal, E., and Garcia-López, M.-À. (2015). Express delivery to the suburbs. Transport Infrastructure and European cities. *ieb.ub.edu*.

Percoco, M. (2016). Highways, Local Economic Structure and Urban Development. *Journal of Economic Geography*, 16(5):1035–1054.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *The Journal of Political Economy*, 90(6):1257–1278.

Schwanen, T., Dieleman, F. M., and Dijst, M. (2001). Travel behaviour in Dutch monocentric and policentric urban systems. *Journal of Transport Geography*, 9(3):173–186.

Schwanen, T., Dieleman, F. M., and Dijst, M. (2004). Policies for urban form and their impact on travel: the Netherlands experience. *Urban Studies*, 41(3):579–603.

Smaal, M. L. (2012). *Politieke strijd om de prijs van automobiliteit: De geschiedenis van een langdurend discours : 1895-2010*. Eburon, Delft.

Veenendaal, A. J. (1996). Railways and the Coming of Road Transport in the Netherlands, 1919-1940. *Business and Economic History*, pages 293–300.

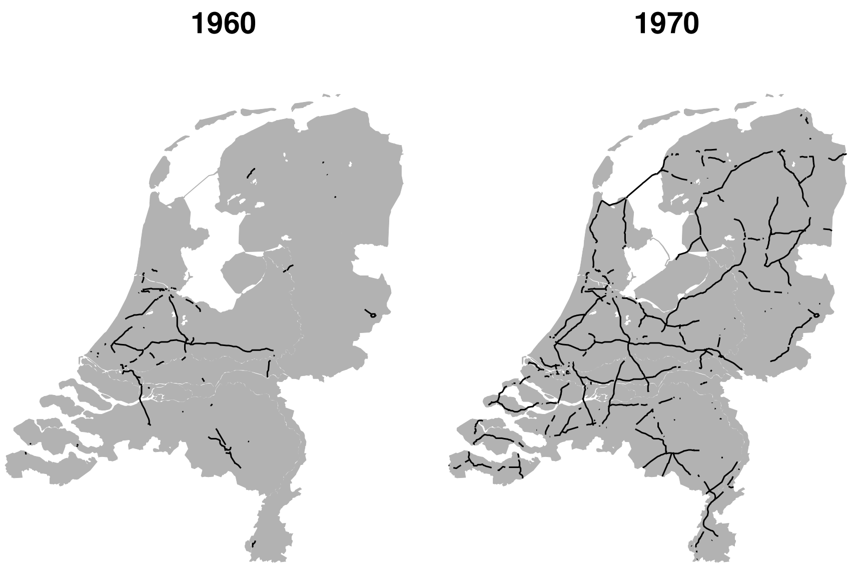
Wooldridge, J. M. (2002). *Econometric Analysis of cross section and panel data.* MIT Press Books. MIT Press.

Appendix A. Road networks

Figure A.1: Main roads in the Netherlands (1821)

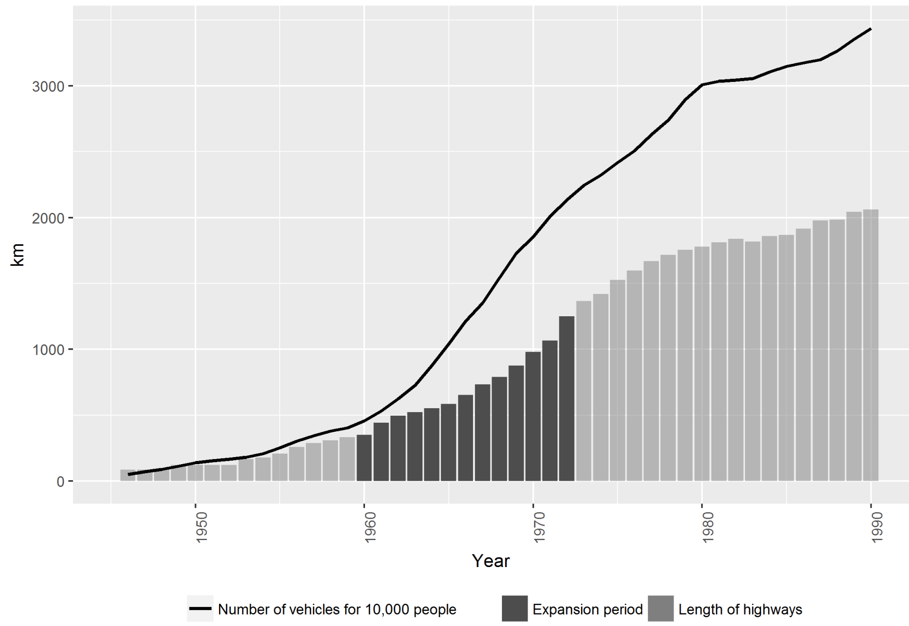
Source: Rijkswaterstaat Beeldbank, multimedia archive (2015) - *Kaart der grote wegen van de 1e klaste met zij kalkvers volgens W.B. 13 maart 1821.* [https://beeldbank.rws.nl/MediaObject/Details/344244 , accessed: December 18th 2016]

Figure A.2: The highway network in 1960 and 1970



Source: Nationaal Historisch Wegenbestand , Data portal of the Dutch government (2015) [https://data.overheid.nl/data/dataset/nationaal-historisch-wegenbestand, accessed: December 19th 2016]

Figure A.3: Highway network length and car ownership (1945–1990)



Source: Ligtermoet (1990), Statistics Netherlands (2016)

Appendix B. Sensitivity analysis

Appendix B.1. Definition of central cities and suburbs

Table B.1: Comparison between different population thresholds for central cities in 1930

Population threshold (1930)	Central cities	Suburban municipalities	Peripheral municipalities
35,000	26	156	629
50,000	20	133	658
65,000	15	116	680

Table B.2: Population growth and highways – Main results under various central cities population thresholds

Dependent variable: Log. pop. growth (1970–1980)

	Pop. Threshold 35,000		Pop. Threshold 50,000		Pop. Threshold 65,000	
	(1)	(2)	(3)	(4)	(5)	(6)
Highway rays*periphery	0.16027***		0.15982***		0.15090**	
	(0.06123)		(0.06184)		(0.05875)	
Highway rays*suburbs	0.02285		−0.00456		0.03348	
	(0.06971)		(0.07982)		(0.07783)	
Highway rays*central city	0.01199		−0.01132		0.02590	
	(0.07360)		(0.10705)		(0.12592)	
Highway density*periphery		0.05676***		0.05877***		0.05553***
		(0.02109)		(0.02155)		(0.02079)
Highway density*suburb		−0.00935		−0.02957		−0.02026
		(0.02374)		(0.02980)		(0.03059)
Highway density*central city		−0.01483		−0.03991		−0.05496*
		(0.02129)		(0.02919)		(0.02946)
Suburbs	0.22993	0.16774*	0.28349	0.21933*	0.22409	0.20796*
	(0.15180)	(0.09290)	(0.18291)	(0.11396)	(0.18224)	(0.11959)
Central city	0.10100	0.00906	0.16184	0.11401	0.08231	0.21315*
	(0.16243)	(0.09848)	(0.26903)	(0.14813)	(0.30149)	(0.12404)
Buffer zone share	−0.15650***	−0.13027***	−0.14900***	−0.12744***	−0.17187***	−0.12693***
	(0.05353)	(0.04044)	(0.05225)	(0.03978)	(0.05241)	(0.03841)
Rail stations (1930)	0.00254	0.00111	0.00238	0.00152	0.00232	0.00070
	(0.00335)	(0.00351)	(0.00345)	(0.00354)	(0.00343)	(0.00357)
Log Pop. (1930)	0.00500	−0.00732	0.00519	−0.00658	0.00423	−0.00826
	(0.03934)	(0.03627)	(0.03940)	(0.03624)	(0.03934)	(0.03620)
Log Pop. (1960)	−0.05486	−0.04862	−0.05521	−0.04869	−0.06010	−0.05089
	(0.04926)	(0.04732)	(0.04862)	(0.04638)	(0.04895)	(0.04616)
Nature coverage share	−0.06976	−0.05071	−0.06932	−0.04736	−0.07692*	−0.04937
	(0.04480)	(0.04847)	(0.04530)	(0.05008)	(0.04501)	(0.04986)
Green heart share	−0.04220	0.03208	−0.03986	0.02458	−0.04325	0.02610
	(0.04012)	(0.02247)	(0.03806)	(0.02279)	(0.03806)	(0.02248)
Distance from central city	−0.00158	−0.04670***	−0.00125	−0.05355***	0.00365	−0.04365***
	(0.01595)	(0.01344)	(0.01876)	(0.01374)	(0.01633)	(0.01409)
Reclaimed land share	0.60774*	0.52549	0.60556*	0.52545	0.59523*	0.51264
	(0.34038)	(0.32002)	(0.34107)	(0.32010)	(0.33949)	(0.31943)
Constant	0.41488***	0.66496***	0.41606***	0.68340***	0.45961***	0.69286***
	(0.09461)	(0.13886)	(0.09289)	(0.13068)	(0.08792)	(0.13155)
Instrumented	Yes	Yes	Yes	Yes	Yes	Yes
Kleibergen-Paap F-statistic	11.54	37.47	10.96	33.62	12.25	34.65
Observations	811	811	811	811	811	811

Notes: (i) Highway variables are in 1970 levels. (ii) Highway density is in logarithm. (iii) Possible endogeneity in the interaction terms is addressed following the single first-stage approach. (iv) Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01

Appendix B.2. Peripheral municipalities subsample

Table B.3: Population growth and highways – Peripheral municipalities subsample

(Dependent variable: Log. pop. growth (1970–1980)				
	IV (1)	IV (2)	OLS (3)	OLS (4)
Highway rays	0.08237**		0.01385***	
	(0.03812)		(0.00390)	
Highway density		0.02868**		0.01052***
		(0.01327)		(0.00271)
Buffer zone share	−0.10029	−0.05930	−0.06766	−0.06042
	(0.06120)	(0.05397)	(0.05381)	(0.05408)
Rail stations	0.00858**	0.00597	0.00812*	0.00727*
	(0.00421)	(0.00419)	(0.00427)	(0.00414)
Log Pop. (1930)	−0.02938	−0.04091	−0.04660*	−0.04672*
	(0.02905)	(0.02717)	(0.02676)	(0.02606)
Log Pop. (1960)	−0.00113	0.00612	0.02331	0.02013
	(0.03358)	(0.03178)	(0.02847)	(0.02795)
Nature coverage share	−0.00194	−0.00084	−0.01462	−0.01119
	(0.05129)	(0.05145)	(0.04924)	(0.04894)
Green heart share	−0.01882	0.03551	0.02775	0.03656
	(0.03720)	(0.02395)	(0.02350)	(0.02353)
Distance from central city	−0.03638***	−0.05444***	−0.04820***	−0.05200***
	(0.01340)	(0.01315)	(0.01278)	(0.01286)
Reclaimed land share	0.37810	0.27595	0.27484	0.26202
	(0.29637)	(0.28204)	(0.27953)	(0.27658)
Constant	0.42688***	0.55136***	0.47355***	0.50807***
	(0.09082)	(0.10886)	(0.09552)	(0.09414)
Instrumented	Yes	Yes	No	No
Kleibergen-Paap F-statistic	10.04	23.79		
Observations	658	658	658	658
R^2			0.1298	0.1365

Notes: (i) Highway extent variables are in 1970 levels. (ii) Highway density is expressed in logarithm. (iii) Robust standard errors in parentheses. *p<0.1, **p<0.05, ***p<0.01

Appendix B.3. Effects on population growth after 1980

Table B.4: Population growth and highways (1980–1990)

Dependent variable: Log. pop. growth (1980–1990)				
	IV (Full sample)		IV (Periphery subsample)	
	(1)	(2)	(3)	(4)
Highway rays*periphery	0.01473		0.04492	
	(0.02163)		(0.02806)	
Highway rays*suburbs	0.04474			
	(0.06457)			
Highway rays*central city	0.08919			
	(0.12148)			
Highway density*periphery		0.01106		0.01958*
		(0.00987)		(0.01187)
Highway density*suburb		0.01255		
		(0.02638)		
Highway density*central city		−0.00480		
		(0.01965)		
Suburbs	−0.08838	−0.02190		
	(0.14764)	(0.08198)		
Central city	−0.35291	−0.02221		
	(0.40563)	(0.10662)		
Buffer zone share	−0.13296**	−0.11177**	−0.28818**	−0.22665**
	(0.06338)	(0.04480)	(0.12102)	(0.08912)
Nature coverage share	−0.02226	−0.02535	0.00891	0.00831
	(0.05189)	(0.04768)	(0.05516)	(0.05171)
Green heart share	0.00103	0.00553	−0.01056	−0.00665
	(0.02246)	(0.02144)	(0.02302)	(0.02037)
Constant	0.54533**	0.58856**	0.56563**	0.70349**
	(0.27458)	(0.28355)	(0.26958)	(0.31864)
Control variables	Yes	Yes	Yes	Yes
Instrumented	Yes	Yes	Yes	Yes
Kleibergen-Paap F-statistic	18.34	41.05	17.68	36.05
Observations	672	672	512	512

Notes: (i) Included control variables are identical to the variables included in Tables 4 and B.3. (ii) Unit of analysis is municipalities in 1990. (iii) Highway extent variables are in 1980 levels. (iv) Highway density is expressed in logarithm. (v) Possible endogeneity in the interaction terms is addressed following the single first-stage approach. (vi) Robust standard errors in parentheses. *$p<0.1$, **$p<0.05$, ***$p<0.01$

Appendix C. First stage regression results

Table C.5: Population growth and highways – Main results (first stage regressions)

	Main mode		Restricted specification	
Dependent variable:	Highway rays	Highway density	Highway rays	Highway density
	(1)	(2)	(3)	(4)
Road density	0.07009***	0.21781***	0.06987***	0.21516***
	(0.02117)	(0.03756)	(0.02109)	(0.03774)
Suburbs	0.09954	0.25874	0.09270	0.17375
	(0.18834)	(0.28181)	(0.18004)	(0.27025)
Central city	0.19029	1.38634*	0.09270	0.17375
	(0.52512)	(0.74724)	(0.18004)	(0.27025)
Buffer zone share	0.59057**	0.59485	0.58723**	0.55334
	(0.24998)	(0.43317)	(0.25015)	(0.42990)
Rail stations (1930)	−0.00130	0.06778	−0.00029	0.08026*
	(0.03073)	(0.04242)	(0.03087)	(0.04188)
Log. pop. (1930)	−0.21466**	−0.31256*	−0.21241***	−0.28463*
	(0.08316)	(0.16854)	(0.08197)	(0.15868)
Log. pop. (1960)	0.32202***	0.71421***	0.32359***	0.73374***
	(0.09229)	(0.16072)	(0.09053)	(0.15408)
Nature coverage share	−0.01797	−0.39332	−0.02032	−0.42251
	(0.34075)	(0.59755)	(0.33960)	(0.59908)
Green heart share	0.59021***	0.30454	0.58710***	0.26595
	(0.17269)	(0.23425)	(0.17428)	(0.23166)
Distance from Central city	−0.22875**	0.27024	−0.24055***	0.12359
	(0.11282)	(0.17408)	(0.08533)	(0.13320)
Reclaimed land share	−1.04179**	−0.83441	−1.02144**	−0.58159
	(0.50954)	(1.01588)	(0.49593)	(0.99425)
Constant	0.74610	−2.72208***	0.75110	−2.65994***
	(0.50657)	(0.88250)	(0.50785)	(0.87998)
R^2	0.096	0.127	0.097	0.125
Kleibergen-Paap F-statistic	10.96	33.62	10.97	32.50
Observations	811	811	811	811

Notes: (i) Highway extent variables are in 1970 levels. (ii) Highway density (1821, 1970) variables are expressed in logarithm, (iii) Robust standard errors in parentheses. $p<0.1$, **$p<0.05$, ***$p<0.01$

Appendix D. Green heart, buffer zones and nature coverage

Figure D.4: Green heart, buffer zones and nature coverage